DOCUMENT RESUME

ED 345 121 CG 024 186

AUTHOR Haffey, Daniel Lee
TITLE Moral Development: A Critical Review of Research on Kohlbergian Stage Theory.
PUB DATE Aug 91
NOTE 95p.; Doctoral research paper, Biola University.
PUB TYPE Dissertations/Theses - Undetermined (040)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Developmental Psychology; Developmental Stages; Epistemology; *Moral Development; Piagetian Theory; Validity
IDENTIFIERS *Kohlberg (Lawrence)

ABSTRACT

Research in moral development from a cognitive-developmental model has been greatly influenced by the Piagetian stage theory of Lawrence Kohlberg. Based on the assumption of inherent cognitive development of internal structures, Kohlberg's theory maintains that the stages are universal across cultures and persons. As a primarily organismic theory, minimal importance is given to external factors such as culture, socialization processes, and situational context. Individuals are expected to progress sequentially through each stage regardless of cultural, personal, or environmental influences. The primary instrument used in assessing moral development in the literature on Kohlberg's theory is the Moral Judgment Interview (MJI) originally developed in Kohlberg's (1958) doctoral dissertation. In general, the empirical evidence provides little support for Kohlberg's assertion of a strict stage concept. Despite a moderately high level of consistency using the MJI, which may be related to instrument bias, the longitudinal data exhibit significant inconsistencies that cannot sustain the "structured whole" assumption. There appears to be significant support for upward sequentiality through Kohlberg's stages, with qualifications. First, the strength of this evidence rests on extensive revisions to the scoring systems which have eliminated deviant findings of regression and stage-skipping. Second, the evidence that does support the notion that the upward sequentiality assumption is limited to the lower stages of Kohlberg's hierarchy. Kohlberg's research program has encountered substantial problems in supporting the empirical existence of principled Post-Conventional reasoning. (LLL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED345121

MORAL DEVELOPMENT: A CRITICAL REVIEW OF RESEARCH ON KOHLBERGIAN STAGE THEORY

A Doctoral Research Paper

Presented to

the Faculty of the Rosemead School of Psychology

Biola University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Psychology

CG024186

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

by

Daniel Lee Haffey

August 1991

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Daniel L. Haffey

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

MORAL DEVELOPMENT: A CRITICAL REVIEW OF RESEARCH ON KOHLBERGIAN STAGE THEORY

by

Daniel Lee Haffey

APPROVED:

[signature] Date 4/13/92
First Reader

[signature] Date 4/13/92
Second Reader

APPROVED:

[signature]
Dean

4/13/92
Date

Copyright © 1992 by Daniel Lee Haffey

ABSTRACT

MORAL DEVELOPMENT: A CRITICAL REVIEW OF RESEARCH ON KOHLBERGIAN STAGE THEORY

by

Daniel Lee Haffey

Research in moral development from a cognitive-developmental model has been greatly influenced by the Piagetian stage theory of Lawrence Kohlberg. The theoretical assumptions of Kohlberg's stage theory are evaluated in terms of the empirical research. These assumptions are found to lack a clear empirical foundation. A complex stage theory which integrates social and individual factors was found to better fit the data than Kohlberg's "simple" stage theory.

TABLE OF CONTENTS

PAGE

ACKNOWLEDGMENTS.....vi

DOCTORAL RESEARCH PAPER

Introduction.....1

Methodological Considerations.....7

Assessment of Moral Development.....8

Characteristics of the MJI.....9

Reliability.....9

Validity.....13

Conclusions.....22

Upward Progression.....24

Support for Upward Progression.....24

Evidence Against Upward Progression.....29

Conclusions.....33

Invariant Sequentiality. 36

Support for Invariant Sequentiality.....37

Evidence Against Invariant Sequentiality.....43

Conclusions.....47

Single Stage Dominance.....48

Evidence Supporting Single Stage Dominance.....49

Evidence Against Single Stage Dominance.....52

Conclusions.....62

Universality.....63

Cultural Differences.....64

Sex Differences.....68

PAGE

Conclusions.....71

Clinical Implications.....75

Directions for Future Research.....76

REFERENCES.....78

ACKNOWLEDGMENTS

Through the support and influence of many people, this paper represents the culmination of my graduate education. I am deeply grateful for the significant contributions to my growth by my professors at Rosemead and particularly for the help and support of my readers, Dr. Pat Pike and Dr. John Kelley. Their theoretical critiques, organizational suggestions and editing have been invaluable in the process of writing this paper. I am also indebted to Dr. Bill Hunter for his expert editing. I must also credit my wife for her consistent encouragement, support and computer skills. Above all, I must thank Dave Cabush for his consistent support of my growth from the beginning. This paper is dedicated to my father, who would have been proud.

MORAL DEVELOPMENT: A CRITICAL REVIEW OF RESEARCH ON KOHLBERGIAN STAGE THEORY

Introduction

Moral development research has been largely dominated by the cognitive-developmental approach in the last two decades. Lawrence Kohlberg has been at the forefront of this field. This paper is a review of the empirical evidence on the stage theory of moral development formulated by Lawrence Kohlberg (1958) in his doctoral dissertation and revised and elaborated upon in subsequent writings (Kohlberg, 1963, 1969, 1971a, 1973, 1978, 1981, 1984). Kohlberg's theory was an extension of Piaget's (1932/1965) work on moral development which held that individuals progress through two stages from heteronomous reasoning to autonomous reasoning. Similarly Kohlberg held that there are progressive, sequential, and universal stages in the development of moral reasoning. Stages were thought to derive from distinct cognitive structures that center on an orientation toward justice as the essence of moral reasoning.

The first level (I) was termed Preconventional and encompassed Stages 1 and 2. In these stages moral reasoning is based on obedience of authority out of a fear of physical punishment (Stage 1) and personal gain from a

fair act of reciprocity with others (Stage 2). The Preconventional individual reasons that the consequences of behavior and the reciprocal gain from acting "morally" are the primary factors in moral decisions. Morality is conceptualized in instrumental terms based primarily on tangible rewards and punishments.

The second level in Kohlberg's theory was the Conventional level (II) which covered Stages 3 and 4. Moral reasoning at the Conventional level is primarily involved with conformity to the expectations of the social order (Stage 3) and demonstrating a duty to the maintenance of societal norms and institutions (Stage 4). Obedience and the upholding of the existing structure of law and order is a central component of this level of moral reasoning.

The final level, the Postconventional level (III), included Stages 5 and 6 where abstract principles of morality govern moral reasoning. Representing the highest strata of moral thought, Postconventional thinkers are concerned with the accepted social contract, chosen from a number of alternatives (Stage 5). Progress to universal objective principles of morality that exist free of societal constraints defines the highest stage of moral development in Kohlberg's theory (Stage 6). Postconventional morality is characterized by internalized principles that are inherent in the human condition,

derived from a respect for the individual dignity of each person. This level of moral thought is held to be possible only after the capacity for abstract thought is obtained in the adolescent years. The developmental thrust in Kohlberg's stage theory is the expansion of the scope of moral reasoning from an egocentric position in Level I to inclusion of society-at-large at Level II, then to a principled morality at Level III.

In current formulations, Kohlberg Stage 6 was dropped except as a theoretical construct due to its lack of empirical support in longitudinal studies (Colby, Kohlberg, Gibbs & Lieberman, 1983; Kohlberg, 1978). Material previously scored at Stage 6 is now considered to represent the B substage at a lower stage such as 5B.

In recent years, Kohlberg and his associates have postulated a "soft" seventh stage, a stage of cosmic oneness and unity of being based on ethical and religious awareness (Kohlberg, 1981; Kohlberg, Levine, & Hewer, 1983). This hypothesized seventh stage has addressed existential issues of meaning, metaphysical concepts of God, and religious epistemologies. Conceptualized as broader in scope than the "hard" stages that are based on a justice orientation, this ultimate level of morality has had no empirical data to support its existence.

According to Kohlberg, stages of moral development are invariably sequential in nature, qualitatively distinct

from each prior stage with each representing a higher level of moral development than the previous level. Kohlberg's theory holds that as the individual develops, the moral understanding of the current level is integrated into the broader and more sophisticated schema of their new stage, thus displacing the previous stage. Conflict between one's current stage and the next higher stage is postulated to be the mechanism that "pulls" the individual toward a more sophisticated and differentiated level of moral reasoning. This mechanism of disequilibrium resulting from stage conflict has its roots in Piaget's (1983) concept of equilibration. Thus, Kohlberg's stage theory held that the development of moral thought is derived from cognitive structures that unfold in response to interaction with the environment to produce increasingly higher, more adequate levels of moral reasoning.

Kohlberg's theory also assumed that the organization of moral reasoning will be consistent across differing situations forming a "structured whole." Moral reasoning will center on the individual's dominant stage with some use of adjacent stages. Such "stage mixture" represents a transition to the next higher stage.

Based on the assumption of inherent cognitive development of internal structures, Kohlberg's theory maintains that the stages are universal across cultures and persons. As a primarily organismic theory, minimal

importance is given to external factors such as culture, socialization processes and situational context. Individuals are expected to progress sequentially through each stage regardless of cultural, personal, or environmental influences.

Kohlberg's stage theory represents the cognitive-developmental approach to moral development, emphasizing the development of cognitive structures as the primary factors in the acquisition of moral thought. Psychoanalytic theory and social learning theory provide alternate views of the acquisition of moral reasoning.

Psychoanalytic theory has attributed morality to the identification and internalization of parental values and prohibitions. Internal irrational forces, primarily guilt, are then responsible for moral behavior. In contrast, the social learning perspective has placed an emphasis on social context and other external factors that influence moral decisions (Bandura, 1977). Social learning theory has viewed moral development as a complex, reciprocal, multidimensional process that cannot be isolated from situational variables by positing uniform mental phenomena such as stages. These variables include the modeling of significant others, personality of the participants and models, social desirability, the social setting, and evaluation of long and short-term consequences. Social learning theorists reject Kohlberg's assertion of an

inherent evolution of moral thinking as reductionistic and rigid, minimizing important individual differences. Empirical research has been conducted by those holding a social learning perspective and has demonstrated the influence of social context and modeling on moral judgments and behavior (Bandura, 1969; Bandura & McDonald, 1963; Cowen, Langer, Heavenrich, & Nathanson, 1969; Prentice, 1972). The focus of this paper is limited to the cognitive-developmental approach as conceptualized by Kohlberg.

Kohlberg's stage theory has generated a considerable body of writing much of which is critical of his theoretical and empirical formulations (Conroy & Burton, 1980; Locke, 1979; Peters, 1971; Simpson, 1974; Sullivan, 1977). The majority of the literature is devoted to theoretical and philosophical arguments about the nature of morality and moral education (Haan, 1982; Leming, 1983; Modgil & Modgil, 1986).

Despite the fact that the bulk of writing on moral development has consisted of a theoretical debate, substantial research has been conducted to test the validity of a stage theory. A significant portion of this has been done by Kohlberg and his associates who are often referred to as the "Harvard Group." A review of the research on Kohlberg's moral development theory has been published (Kurtines & Greif, 1974) but is dated. Given

that there have been a significant number of empirical studies completed since this review, there is a need in the literature for a review of the current status of the empirical research on Kohlberg's stage theory of moral development.

The purpose of this paper is to systematically review the empirical evidence on Kohlberg's theory of moral development. Methodology used in the assessment of moral development is the initial consideration as it is an essential connection between theory and the empirical literature. In the following sections, the review of the empirical literature is organized around four primary underlying assumptions of Kohlberg's stage theory: (a) the progressive upward movement through the stages, (b) the invariant sequential ascension through the stages, (c) the consistent use of a single dominant stage of moral reasoning (structure), and (d) the universality of moral stages. A summary and conclusions section based on the empirical review is followed by a discussion of the clinical implications of Kohlberg's stage theory and directions for future research.

Methodological Considerations

Integral to an investigation of the evidence for a stage theory of moral development is an understanding of how the individual's moral reasoning is operationalized in

terms of a specific stage. Bridging the gap from subjective theory to objective measurement is an inherent methodological problem in moral development research and is a focal point in the debate over the validity of Kohlberg's stage theory. This section is a discussion of the methodological factors in moral development research.

Assessment of Moral Development

The primary instrument used in assessing moral development in the literature on Kohlberg's theory is the Moral Judgement Interview (MJI) originally developed in Kohlberg's (1958) doctoral dissertation. The scale and the scoring system have undergone several revisions from the original Ideal Type (Kohlberg, 1958) to the Structural Issue System (Kohlberg, 1971b) then to the Standardized Scoring System (Kohlberg, Colby, Gibbs & Speicher-Dubin, 1976) and finally to the Standard Issue Moral Judgment Interview and Standard Issue Scoring System (Colby & Kohlberg, 1987; Colby, Kohlberg, Gibbs, Candee, et al., 1983).

The basic format of the current form of the MJI, the Standard Issue Scoring System, uses the MJI to present three hypothetical moral dilemmas. Each dilemma focuses on a conflict between two central moral issues such as preserving life and upholding the law. The subjects are asked to make a moral judgment regarding the moral conflict. The interviewer then asks 9 to 12 "probe"

questions to elicit the reasoning behind the given answer. Responses are coded by matching them with seven criterion cases in the scoring manual.

The Standard Issue system has three parallel forms, A, B, and C, containing three dilemmas each. The Standard Issue Scoring manual, practice cases, and tutorial have been recently published and provide a detailed look at the scoring system (Colby, et al., 1983; Colby & Kohlberg, 1987).

Characteristics of the MJI

Reliability. The first reliability data on the Moral Judgment Scale were published by Kohlberg (1958) and expanded by Kurtines and Greif (1974) to include other studies. Bergling (1981) also cited reliability statistics from additional studies. Most recently Colby et al. (1983) have provided reliability figures from a 20 year longitudinal study based on Kohlberg's (1958) original work.

Empirical research on Kohlberg's stages of moral development has evidenced some general problems in the reliability and validity of the stage criteria and scoring (Kurtines & Greif, 1974). The criticisms from Kurtines and Greif's (1974) review fall into three areas. First, they were critical of the lack of standardization of the interview and coding system. Most of the studies they reviewed used different dilemmas and coding criteria making

comparison across studies difficult. Second, the reliability of the coding system was questioned, particularly the lack of published data on test-retest and parallel forms reliability. Data on internal consistency were also missing from the studies reviewed by Kurtines and Greif. In addition, the MJI is criticized as time-consuming, subjective, and unavailable for replication studies.

The longitudinal data published by Holstein (1976), Kuhn (1976), and Kohlberg and Kramer (1969) have also shown discrepancies in the sequential stage hypothesis related to problems in defining the stages or in the reliability and validity of the scale and scoring system.

Kuhn's (1976) short-term longitudinal study was unable to eliminate measurement error as an explanation for short-term regressions. Kuhn suggested that instruments assessing moral judgments inherently contain more measurement error than other types of assessment instruments. Holstein (1976) also found significant regressions in her 3 year longitudinal study but could not rule out measurement error as a possible explanation for these results.

Perhaps most significant are the difficulties encountered in Kohlberg and Kramer's (1969) longitudinal study with "sophomore" retrogressions, the Stage "4 1/2" phenomenon. Subjects who scored earlier at Stages 4 and 5

regressed on retests to Stage 2 between high school and their second year of college. Rather than being seen as invalidating a stage theory these problems resulted in the development of a revision of the scoring system. In 1971 Kohlberg (1971b) developed the Structural Issue Scoring system which relied on an expanded "sociomoral perspective" in an attempt to address these problems. However, problems with the scoring system continued largely due to overly general and abstract scoring definitions and criteria.

Partly as a response to these continued problems, the Standard Issue Scoring System (Colby & Kohlberg, 1987; Colby et al., 1983) was developed with the intent to "achieve greater objectivity and reliability in scoring by specifying clear and concrete stage criteria" (Colby et al., 1983, p. 8). In scoring the MJI, the primary change from previous systems involved the refining of the unit of analysis toward a more objective and conceptually coherent unit which is then scored by matching to specific and concrete criteria found in the scoring manual. Reliability studies conducted by Colby et al. (1983) indicated that significant improvements in reliability have been achieved through this revision of the scoring system.

Test-retest reliability studies using the Standardized Scoring System (Colby et al., 1983), indicated reliability on Forms A and B were well within accepted limits. No data were available for Form C.

Interrater reliability studies (Colby et al., 1983) using scales ranging from 9 to 13 points, showed agreement from 88% to 100% within 1/3 of a stage, 75% to 88% for the 9-point scale, and 53% to 63% for the 13-point scale. Interrater reliability decreased somewhat as increasingly discriminating scales were used (100%, 75%, 67%).

The results of alternate form reliability studies have been consistent with the other reliability studies. Correlations between alternate forms range from .84 to .95.

The reliability data have shown that the revision by Colby et al. (1983) has adequately addressed the criticisms of Kurtines and Greif (1974) concerning the lack of standardization, as well as the reliability and the validity of the coding schemes. These changes enhanced the psychometric properties of the MJI and allowed greater objectivity in assessing moral judgments. However, the reliability data on the revised Standard Issue Scoring System have not yet been replicated by an outside source. Additional longitudinal studies are needed to determine the consistent reliability of the Standard Issue Scoring, particularly with populations other than the white middle-class sample employed by Kohlberg.

Kuhn (1976) argued that the MJI should not be evaluated according to traditional psychometric methods, such as relationship to external criterion (moral action), for assessing reliability. Her argument was based on

Kohlberg's theoretical emphasis on the sequentiality of moral stages rather than the reliability of classification by stages. Validation of the sequence assumption infers support for the reliability and construct validity of the assessment instrument. Kuhn stated that "it is impossible that repeated testing could show individuals to progress through the stages in an orderly manner if their assignment to stages were not reliable" (p. 163). Colby et al. (1983) concurred with Kuhn that construct validity (i.e., confirmation of the stage theory hypothesis), not predictive validity, is the only appropriate test of the instrument's reliability.

Validity. A study by Rubin and Trotter (1977) indicated that the type of format, verbal or multiple choice, has a significant effect on moral judgment scores. The use of a multiple choice format resulted in significantly higher scores over the verbal format, suggesting that verbal interviewing as used by Kohlberg may underestimate the level of moral judgment. That is, Kohlberg's interview may be an assessment of verbal abilities as well as moral reasoning, thus confounding verbal sophistication and moral stage. Rosenthal and Zimmerman (1978) asked "Can inability to verbalize elegantly be taken as a failure to comprehend the chosen rationales, or is this merely a reflection of verbal skills, distinct from understanding?" (p. 165). Rest

(1976) found similar results suggesting that moral judgment may be underestimated when assessed verbally, a problem that Rest (1976, 1979, 1986) sought to remedy in his nonverbal Defining Issues Test (DIT). Rest found that subjects scored almost two stages higher on the DIT than on Kohlberg's verbal measure. Differences based on response mode are problematic to the validity of the MJI (Rest, 1979).

In a 1976 study, Levine found that using real persons (mothers or close friends) rather than fictitious characters in the MJI's dilemmas produced significantly different scores, that is, Stage 3 responses significantly increased and Stage 4 responses significantly decreased when primary others are used. This discrepancy may be related to less affective involvement in fictitious protagonists than with real-life characters. Leming (1975) also found a significant difference in dilemmas that centered on practical situations than hypothetical moral dilemmas. These results suggested that the MJI may be biased toward producing higher stages of moral reasoning than are representative of the subject's real-life moral decisions. This discrepancy may be related to the difference in affective involvement between hypothetical and relevant dilemmas. The unclear connection between hypothetical dilemmas and real-life situations limit the generalizability and validity of the MJI.

Evans (1986) raised the question of whether studies using Kohlberg's interview technique and employing a pre-test and post-test design are confounded by a practice effect. Incidental learning of moral dilemmas is a significant factor that may impact reliability by overestimating moral stages. Bergling (1981) also noted that Kohlberg has not discussed the possible effect of repeated exposure to the same dilemmas in his longitudinal data (Colby et al., 1983). The neglect of learning factors in the MJI may be related to the cognitive-developmental assumption that cognitive structures take precedence over learning. Such inattention to possible confounding factors, however, clouds the validity and reliability of the assessment measure.

Simpson (1974) alleged that the MJI is ethnocentric, that it does not take into consideration cultural differences and is not valid for use with non-Western cultures. Simpson pointed to language and value differences that bias moral reasoning outcomes toward lower stages in non-Western cultures, a charge that was denied by Kohlberg (Kohlberg & Kramer, 1969; Kohlberg, Levine, & Hewer, 1983). Simpson argued that the relationship between verbal skills required by the MJI and higher stages of moral reasoning prevents an accurate assessment of cultures where education does not keep pace with Western cultures.

Kurtines and Greif (1974) also found the MJI lacking in predictive validity, asserting that the relationship between moral reasoning and moral action was not demonstrated in Kohlberg's research. For support they cited a study by Haan, Smith, and Block (1968) that compared the moral development of political activists to that of nonactivists. The results showed that 75% of Stage 6 men participated in a free speech sit-in but 60% of the Stage 2 men also participated. The results further indicated a discrepancy in the expected relationship between moral stages and moral action as Preconventional subjects joined principled subjects in a demonstration for free speech.

This apparent anomaly was shown to not be an indictment of the predictive validity of Kohlberg (1983) as rescoring using the revised scoring system showed no subjects scoring at Stage 2 as originally reported by the Haan et al. (1968) study. The apparent Stage 2 protesters were actually in transition to Stage 5, adopting a radical ideology in rejection of Conventional reasoning (Kohlberg & Candee, 1984).

Kohlberg and Candee (1984) advanced the idea that moral stage and moral action evidence a monotonic relationship, that the higher the moral stage the more likely that moral action will be consistent with the reasoning at the person's dominant stage. The authors contended that moral

stage influences moral action in two ways: through (a) a deontic choice of moral rightness and (b) a sense of responsibility to act on that judgment. A re-analysis of the data from the Haan et al. (1968) study on the Free Speech movement and the Milgram study (Kohlberg, 1969) are cited as a demonstration of the monotonic relationship between moral stage and moral action.

Blasi (1980), in a review of studies linking moral judgment and moral action, found problems in defining this relationship while also finding qualified statistical support for a cognitive-developmental perspective on the structure-to-behavior connection. Looking at the link between moral stage and other variables such as delinquency, honesty, altruism, and conformity, Blasi concluded that there is "considerable support" for asserting that moral reasoning and moral action are statistically related. The strength of this relationship, however, varies from one content area to another. Blasi noted that there are significant problems in the conceptualization and operationalization of the criteria for moral reasoning and behavior. These flaws limit the conclusions that can be made from current research. The author suggested that "One may reasonably expect to find higher and more consistently significant statistical correlations, once research is designed and executed with more care" (Blasi, 1980, p. 37).

Additional research to clarify the functional relationship between the level of moral reasoning and behavior is clearly needed (Kudnick, 1986). The value of defining the parameters of this relationship is substantial for clinical applications and moral development research.

The psychometric issue of regression toward the mean in Kohlberg's scale and a ceiling effect have also surfaced in the research using Kohlberg's assessment methods. Several studies have demonstrated a tendency for more change in lower moral stages than higher stages when initial baseline stages are lower in Kohlberg's scale (Arbuthnot, 1975; Holstein, 1976; Keasey, 1973; Tracy & Cross, 1973). Movement toward a modal point in Level II means that individuals initially rated at Level I (Stages 1 & 2) are more likely to shift to Level II (Stages 3 & 4) than those who began at Level II are to move into Level III (Stages 5 & 6). This differential in upward change may be due to a statistical regression toward Level II mean or may indicate a "ceiling effect" where upward movement is a function of available room for change.

The study by Tracy and Cross (1973) demonstrated this apparent measurement problem in Kohlberg's scale. Their study included 76 seventh grade boys in an effort to determine the effective variables in shifting the subjects' moral stages upward to higher levels. The results of their study showed that Level I (Preconventional) subjects

increased their scores significantly more than Level II (Conventional) subjects regardless of condition. The Level I control group evidenced more change than the Level II experimental group. The authors hypothesized that the nature of Kohlberg's scale allows more room for change at lower levels than at higher levels, demonstrating a "ceiling effect" at the upper stages relative to lower stages. This upper limitation is significant in that most individuals do not ever attain the higher stages, in fact, Stage 6 has no significant empirical support.

This tendency for greater change at lower levels may be related to the cognitive development of the children. Kohlberg held that the level of formal operations is prerequisite for Level III moral judgments, a level that the subjects would not have yet attained (Flavell, 1985). Another possibility is that children at lower stages are more susceptible to social influences (the pre-test interview) than higher level children. However, the results of Tracy and Cross's (1973) study seemed to point to an unevenness in Kohlberg's scale.

Arbuthnot's (1975) study, which used role-playing to induce shifts in levels of moral reasoning, also supported a regressive bias toward the mean. In experimental conditions where confederates role-played a stage below the subject's pre-test stage, the results indicated a backward slide toward the next lower stage. Subjects in Stages 2

and 3 showed the greatest amount of change. Subjects in Stage 4 showed little change. Those in Stage 5 regressed while Stage 6 subjects showed no change. Similar to Tracy and Cross's (1973) findings, more overall change was found at lower levels in Kohlberg's stages than at higher levels.

These findings suggested an inherent psychometric problem in Kohlberg's scale of a statistical regression toward the overall mean in the Conventional level. The research indicated that Kohlberg's method of assessing moral development in terms of discrete stages is not uniform and is subject to the pull of statistical regression and the limitations of a ceiling effect. The cause of this measurement problem may be complex, involving cognitive, social, and theoretical variables.

The introduction of the revised scoring system, Standard Issue Scoring (Colby et al., 1983), which redefined the scoring criteria for upper level stages, appears to have addressed these measurement problems. Only additional research using the new scoring system will provide evidence of improved psychometric properties for Kohlberg's scale.

The inherent methodological difficulty with the "recalibration" revisions in the MJI, and Kohlberg's research program in general, is the method of theorizing Kohlberg used in developing his theory called "bootstrapping" or the "abductive" method (Colby, 1978;

Colby & Kohlberg, 1987). Rather than a hypothetico-deductive method that tests hypotheses deductively, a circular loop between theory and research was employed where results of empirical studies are used to refine the assessment procedures to more closely correspond with the theoretical constructs. For example, the anomalies, regressions and stage-skipping, found in a longitudinal examination of Kohlberg's original sample prompted revisions that eventually eliminated these deviations. The trend is toward greater approximation of the underlying theoretical constructs as opposed to an acceptance or refutation of a series of "a priori" hypotheses. As Bergling (1981) noted, such "bootstrapping" does not allow for verification through replication with new data. Thus a stage theory can be only demonstrated, not verified, with a bootstrapped study; outright verification through independent replication is not permitted with this self-contained method. The significance for Kohlberg's use of a "bootstrapping" methodology is the difficulty in comparing results across studies. Direct comparisons are not possible as Kohlberg's research program was focused on successively closer approximations of the postulated underlying construct and not on experimental designs that are open to replication.

An additional and related problem with Kohlberg's research program historically has been the inaccessibility

of scoring manuals and other methodological data for use in replicating Kohlberg's work. Much of such material needed for replication studies has been circulated in unpublished manuscript form. Until recently (1983) the scoring manual for the Moral Judgment Interview was available only from Kohlberg's Center for Moral Education at Harvard University. This partisan cloistering of essential data has worked against independent replications of Kohlberg's research.

Conclusions

Methodology is a central consideration in establishing the empirical support for Kohlberg's stage theory. The research testing Kohlberg's theoretical assumptions and the conclusions that can be drawn from this research are limited by methodological flaws. This section suggests that there are significant methodological problems in the current research.

Despite substantial improvement in the scoring system, there remains considerable debate about the accuracy and validity of the Moral Judgment Interview (Rest, 1986). Cultural and intellectual bias, unclear meaning of numerical scores, regression to the mean, and the numerous revisions of the scale continue to cloud the usefulness of the assessment instruments.

Relating moral judgment to moral behavior is another problematic area in the research which limits the value and

meaning of assessing stages of moral development. Until recent years (Kohlberg & Candee, 1984), Kohlberg and his associates had largely ignored this important area of research claiming that one can only assess moral reasoning, not moral action itself. A stage theory of moral development that cannot demonstrate the connection between stages and moral action is limited and lacking in utility. Additional study of a substantial nature is needed to define and test these relationships.

The major assumptions of Kohlberg's theory rest on the concept of sequentiality, that moral development is progressive and invariant in nature, reflecting an underlying structure. Empirical validation of these assumptions requires a longitudinal design, one that samples development over time to determine the direction and nature of change. Cross-sectional designs are less powerful in that they often confound age with development. That is, age trends per se do not provide direct support for a stage theory although a general correlation would be expected. Although relatively few longitudinal studies are available in the literature the studies that have been conducted are the primary basis of this review and are evaluated in the following sections.

Upward Progression

The theoretical assumption underlying upward progression in Kohlberg's stage theory was that each progressive stage displaces the previous stage resulting in a qualitative change in the individual's cognitive organization. Consistent with this assumption, the empirical evidence should show an invariable pattern of upward movement through the stages without regression to lower stages of moral reasoning. Regression to a lower stage violates the assumption of upward progression. This assumption is related to invariant sequentiality but is more concerned with the direction of development rather than the manner or pattern of development. This section evaluates the empirical support for the assumption of upward progression.

Support for Upward Progression

Kohlberg's own longitudinal research program has encountered regressive phenomena that initially contradicted the assumption of upward progression. A later analysis of the longitudinal data apparently accounted for the original regressions and found considerable support for upward progression (Colby et al., 1983). As the various studies based on Kohlberg's (1958) original longitudinal sample are the primary support for upward progression, the process from initial problems to revisions held to support Kohlberg's theory are outlined here.

Kramer's (1968) longitudinal study on Kohlberg's original 1958 sample showed significant deviations from expected upward movement through the stages. These deviations included regressions to lower levels of moral reasoning in late adolescence. A follow-up reporting of these data (Kohlberg & Kramer, 1969) also showed significant regressions in those initially scoring at higher stages. The most significant example of regression was a slide from Stage 4 to Stage 2 in approximately 20% of the college student sample. After scoring at higher stages (4 & 5) during high school, these subjects demonstrated a dramatic reversal accompanied by radical political views and rebellious behavior. By age 25 these subjects had returned to a mixed Stage 4 and 5. Kohlberg termed these regressions "sophomore retrogression" and attributed them to an inadequate conceptualization of Stages 4 and 5, specifically the lack of clear differentiation between form and content. Kohlberg explained this deviation as a functional rather than a structural regression. Reverting temporarily to lower levels of moral reasoning may serve the function of promoting separation through rebellion related to Erikson's (1963) ideas about the formation of identity in adolescence. Thus, the observed regression is apparently a part of a normal developmental task not a reversal in the underlying structure of moral reasoning (Kohlberg, 1973; Kohlberg & Kramer, 1969).

A Stage 4 1/2 was proposed to account for these anomalies but this change explained relatively few of the regression cases and was eventually discarded (Colby et al., 1983; Kohlberg, 1973) leaving open the question of whether the observed "retrogressions" in a college population can be understood using a stage theory model. Kohlberg's moving away from a cognitive explanation of these regressions to a psychological basis using Erikson's work with adolescent identity formation also seems to suggest something other than a structural component operating in moral judgment in late adolescence.

In a reanalysis and update of the original 1956-1968 longitudinal studies by Kohlberg (1958) and Kramer (1968), Colby et al. (1983) presented substantial support for Kohlberg's cognitive-developmental approach. The data from Kramer's 1968 study were re-evaluated using the newer Structural Scoring System (Colby et al., 1983). Additional data from the same sample group were added to extend the longitudinal study, providing a 20 year perspective on the subjects' moral development.

Colby et al. (1983) used the original 1958 sample of 58 boys, stratified by age, social class, and sociometric status. The subjects were administered Kohlberg's Moral Judgment Interview at 3-4 year intervals. The boys, ages 10, 13, and 16 at the time of the first administration,

were given follow-up testing. A total of 58 subjects completed the sequence of six testing times.

The results of the study showed a low percentage (5%) of reversals across three parallel forms of the Moral Judgment Interview. The authors attributed these figures to measurement error since short-term test-retest reliability studies of the Standard Issue Scoring system evidenced over twice the reversals found in their actual analysis. The test-retest differences were interpreted as originating from measurement error rather than any actual fluctuations in moral development. The authors supported this interpretation by citing equivalent reliability figures for interrater, parallel, and test-retest studies which suggested an inherent error in the scoring system of about 1/3 of a stage.

Of particular interest in Colby et al.'s (1983) reanalysis is the "sophomore retrogression" found in Kohlberg and Kramer's (1969) study. The two deviant cases of "retrogression" from this study are not viewed as questioning the assumptions of the stage model. The authors took the position that "college-age relativism" represents a "metatheoretical position that can be taken at a number of different developmental stages" (1983, p. 72). The introduction of this "metatheoretical position" apparently eliminates "sophomoric regression" as a regressive phenomenon and places it outside Kohlberg's

stage theory. However, Kohlberg's concept of a transitional stage between Stages 4 and 5 has not been found to be a stable and valid phenomenon (Kohlberg, 1984).

The study by Colby et al. (1983) provided the most compelling support for upward progression as well as other theoretical assumptions of a Kohlbergian stage theory. The significance of the evidence from this study stemmed from the power of the longitudinal data and the strong reliability of the revised Standard Issue Scoring system.

Other longitudinal studies have found that change is consistently progressive (Erickson, 1980; Nisan & Kohlberg, 1982; Page, 1981; Snarey, Reimer & Kohlberg, 1985) reporting regressions in modal stage scores ranging from 1% to 10%. These regressions are within the limits of measurement error reported by Colby et al. (1983). Walker (1986) concluded that these studies, using the current scoring system, "strongly support" the upward progression assumption.

Intervention studies which seek to induce upward change experimentally through exposure to higher stages (+1, +2) and lower stages (-1, -2) of reasoning provide some additional support for upward progression. In a classic study, Turiel (1966) provided some support for upward change in a study that exposed subjects to stages that were -1, +1, and +2 from their pre-tested dominant stage. No significant regressions were found in Turiel's study

although small treatment effects and methodological problems limit the strength of these results. In a similarly designed study, Walker (1982) found no regressions with the +1 condition most effective in producing upward movement. Walker's findings were consistent with other intervention studies (Arbuthnot, 1975; Keasey, 1973; Tracey & Cross, 1973) which demonstrated the effectiveness of the +1 treatment in producing upward progression.

Other intervention studies (Blatt & Kohlberg, 1975; Colby, et al., 1977) used group moral discussions that presented a range of stages. These studies determined that exposure to other stages generally led to movement toward a higher stage of moral reasoning, supporting the upward progression assumption. However these studies also showed some regressions in post-tests which the authors attributed to measurement error.

Evidence Against Upward Progression

Other research has uncovered evidence of regressions that seem to contradict the assumption of upward progression. Longitudinal studies by Holstein (1976), Kuhn (1976), White et al. (1978) and Murphy and Gilligan (1980) have shown problems with regressions.

Holstein (1976) used 52 American families in her 3 year longitudinal study. Each family had a 13 year old son or daughter and scored high on SES measures. The family

members were administered five dilemmas at Time 1 and were then retested three years later, following the design used by Kohlberg and Kramer (1969). The responses were scored using Structural Issue Scoring (Kohlberg, 1971b).

In 3 year post-tests, Holstein found a significant tendency to regress in both adults and adolescents who pre-tested at higher stages (Stages 4-6). Regressions at higher stages ranged from 20-33% across age groups while only 0-2% of lower stage subjects regressed. Holstein's results showed that, at lower stages (Preconventional to Conventional) the assumption of upward advancement was supported by the data. However, in higher stages (Conventional to Postconventional) upward progression is not supported, that is, downward regression was found.

As Holstein noted, measurement error cannot be ruled out as a possible explanation for these regressions. Short-term fluctuations cannot be eliminated as a source of these regressions or progressions as there were no short-term controls used to establish measurement error. Despite this design flaw it is equally possible that the observed regressions do actually represent downward fluctuations in the subject's level of moral reasoning. It is also significant to note that the percentage of regressions found by Holstein (20-30%) exceeded the level of measurement error expected for a reliable instrument.

An alternate interpretation offered by Holstein (1976) was that once cognitive prerequisites are met, other specific and personal variables account for the individual's choice to use reasoning below his or her established capacities. This plausible option highlights the need for further investigation into the complex factors that influence moral choices such as situational context, affective involvement and personality variables.

Colby et al. (1983) responded to the Holstein results by suggesting that Holstein's use of the older Structural Issue Scoring system rather than the revised Standard Issue Scoring procedures offered a possible explanation for the regressions. Support for this assertion was drawn from Kramer's 1968 data which showed regressions that were "almost identical" to those found in Holstein's (1976) study. When Kramer's data were reanalyzed using the revised Standard Issue Scoring system virtually all deviations were eliminated. Although progressive revisions in the scoring system are desirable this "retroactive" validation to explain discrepant findings did not allow clear empirical testing through experimental replication of contradictory results.

In a short-term longitudinal study, Kuhn (1976) tested 5 to 8 year olds at 6 month intervals and found evidence of regressions within a 6 month period while also finding significant support for upward movement over the longer 1

year period. At the 6 month retest point, subjects were as likely to have regressed or remain unchanged as to have advanced. Retesting at 1 year yielded significant progressive change with minimal regression (2%). The regressions at 6 months could be attributed alternately to measurement error or temporal changes in the individual's moral judgment superimposed on an overall upward movement. Kuhn's study demonstrated clearly the short-term inconsistencies in the individual subject's stage assignment. It is also important to note that the results of Kuhn's study, as in the study by White et al. (1978), pertained only to stages 1 to 3 and provided no support for upward progression in higher stages.

White et al. (1978) demonstrated some problematic regressions, specific to certain age groups (11-12; 15-16), while providing some overall support for an upward trend with age. The implication of these regressions was that cognitive-developmental stage theory may neglect relevant environmental factors that operate in moral reasoning. The authors noted that social and cultural factors played a significant role in the moral judgment of their Bahamian subjects. Accounting for these factors is particularly important in assessing the moral reasoning of non-Western cultures.

Arbuthnot (1975) found evidence of downward movement in his study using role-playing to induce shifts in levels of

moral reasoning. In experimental conditions where confederates role-played stages above and below the subjects pre-tested stage, the results indicated a backward slide toward the next lower stage in Stage 5 subjects while lower stage (2 & 3) subjects showed the greatest amount of upward change. Stage 6 (n = 3) subjects showed no significant change.

Bergling (1981), in a critical review of the validity of Kohlbergian stage theory, used a sophisticated log-linear computer model to statistically analyze the qualitative data from longitudinal studies on Kohlberg's stages. Based on this analysis, Bergling rejected the hypothesis that there are more progressive than regressive changes in the higher stages of moral reasoning (4-6) while accepting the hypothesis that there are more progressive than regressive changes in the lower stages (1-3) for children and adolescents.

When the lower stages were regrouped to include Stages 1-4 and the higher stages were defined as 5 and 6, the same conclusions were reached based on the statistical evidence. That is, upward progression was sustained for Stages 1-4 in children and adolescents but results were inconsistent for stages 5 and 6 in adults.

Conclusions

The research seems to have given greater support to the assumption of upward progression in the lower stages (1-3)

than in the more advanced stages (4-6) where the delineations of the stages are less clearly defined or may not represent actual structural stages. Related to this, the evidence also seems to have supported more clearly an upward trend in studies using children and adolescents than those using adults, thereby casting doubt on Kohlberg's stage concept of adult development. This evidence is consistent with the isomorphism between the development of logical functions and moral reasoning, suggesting a stronger connection between development of cognitive structures and moral development than is allowed by Kohlberg's theory. That is, the alternate hypothesis that moral development can be explained largely in terms of cognitive maturation rather than stage-dependent structures of moral reasoning has gained support from the lack of clear evidence for moral stages in adolescents and adults. The nature of adult moral development is as yet unclear and is problematic for a Kohlbergian stage theory. It seems reasonable that adult moral development is based on a variety of individual factors and is more complex than Kohlberg's stage theory allows.

Additionally, the strength of the supportive evidence is diluted by the use of revised scoring systems to account for observed anomalies in previous studies. Regressions have been eliminated in a post hoc fashion through reanalysis with a newer scoring system, thereby reducing

the credibility of Kohlberg's stage theory. While the newer scoring system exhibits greater reliability than prior systems it is unclear whether the revisions represent a better assessment or a means of eliminating deviant results.

Methodological problems limit the strength of the conclusions that can be reached from the current research regarding upward progression. Problems in controlling for measurement error through temporal stability studies preclude a clear understanding of the source of observed regressions and fluctuations. Short-term inconsistencies have been demonstrated to affect moral reasoning, thus suggesting strongly that this measurement problem must be clearly defined and controlled before observed regressions can be attributed with confidence to measurement error.

The less than clear support for the assumption of upward progression, particularly at the upper levels, indicates a need for additional research to provide clearer evidence for this central concept. Further work needs to focus on mediating variables that influence the direction of moral judgments such as affective involvement, situational cues and personality parameters. A clearer understanding of the relationship between cognitive development and moral development would be helpful in differentiating these two domains and in explaining the regressions in upper levels of the scale.

Overall the evidence for upward progression is mixed and does not represent conclusive empirical validation of the assumption of upward progression demanded by Kohlberg's stage theory. Problematic regressions are not clearly accounted for by measurement error, particularly considering that one unexplained regression invalidates the stage theory. There does appear to be a developmental trend in the lower stages which provides some support for upward progression. However, a developmental age trend alone cannot provide conclusive support for a stage theory.

The empirical evidence for upward progression is not of the strength that is required of Kohlberg's stage theory assumption of upward progression, particularly in light of the scoring system revisions that have erased apparent regressions.

Invariant Sequentiality

Borrowing from the cognitive-developmental theory of Jean Piaget, Kohlberg's stage theory follows a sequentiality stage paradigm similar to Piaget's (1954) understanding of the developmental sequence of cognitive abilities. It is hypothesized that the development of moral reasoning follows a stepwise progression of stages with each stage representing a qualitatively differentiated form of cognitive organization. Kohlberg's theory holds that individual development follows an unvarying sequential

pattern through each successive stage, that is, skipping stages violates the assumption of sequentiality. An example would be finding a Stage 3 individual who jumps to Stage 5, skipping Stage 4.

Additionally, following the hierarchial integration concept which holds that successive stages are increasingly difficult and cognitively complex, it is expected that the stages show a hierarchial sequential pattern in the individual's comprehension, preference, and usage. That is, a person will prefer stages higher than can be fully comprehended and will comprehend and spontaneously use dominant and lower stages. This section will evaluate the evidence for invariant sequentiality.

Support for Invariant Sequentiality

Longitudinal studies provide the most direct evidence for sequentiality as repeated testing can assess an individual's movement through the stages. The longitudinal data collected by Kohlberg beginning with his doctoral dissertation in 1958 have provided general support for stage sequentiality although there were some significant discrepancies reported in the initial 12 year analysis (Kohlberg & Kramer, 1969; Kramer, 1968). The original analysis of the longitudinal data by Kramer (1968) indicated that 14% of the older adolescent subjects skipped stages.

Colby, et al. (1983), in a 20-year reanalysis and update of Kohlberg and Kramer's longitudinal data using the recently-developed Standard Issue Scoring system (Kohlberg, Colby, Gibbs, Speicher-Dubin, & Power, 1978), appeared to provide substantial support for the assumption of stage sequentiality. The use of the refined scoring system in the Colby et al. study accounted for almost all the anomalies found in Kramer's (1968) original analysis of the data. The results showed that no case on any of the alternate forms reached a stage without going through the previous stage (Colby et al., 1983, pp. 29-35) based on repeated tests at 3-4 year intervals spanning 20 years. The absence of stage-skipping in this extensive longitudinal study appears to lend considerable support to the invariant sequence assumption. This support must be qualified by the lack of stage-skipping resulting from substantial revisions in the scoring system.

Kuhn (1976) conducted a short-term (1 year) longitudinal study with 5 - 8 year olds. The study supported sequentiality overall, that is, there was no evidence of stage-skipping. However, Kuhn's study can offer only limited support for sequentiality as the span of the study was relatively brief and may not have provided adequate time for stage change. An additional limitation is that the study covered only the first three stages in Kohlberg's sequence.

Intervention studies which introduce experimental conditions to induce changes in moral stages provide another line of evidence in assessing the support for sequentiality. In an original intervention study on the sequentiality of stage theory, Turiel (1966) exposed subjects to moral stages above (+1 & +2) and below (-1) their own dominant stage as determined by a pretest using Kohlberg's moral dilemmas. Turiel expected to find that the +1 condition would be more effective in producing upward change than the other experimental conditions. The results of Turiel's study showed that a +1 stage condition was indeed the most effective in changing moral reasoning while the +2 treatment condition was the least effective. Thus Turiel's study appears to confirm the hypothesis of sequential progression. However the statistical significance of Turiel's results is relatively small. Additionally, the moral judgment scores of the control group regressed as much as the +1 treatment group suggesting significant instability and equivocal results. Turiel's results lend some limited support for sequentiality but does not allow strong conclusions to be drawn (Kurtines & Grief, 1974).

More recent intervention studies by Keasey (1973) and Walker (1982) provide experimental support for sequentiality. Keasey used models to expose subjects to the three experimental conditions: opinion without

supportive reasoning, opinion with reasoning, and reasoning at the +1 stage from the subject's dominant stage as determined by a pretest. Keasey found that exposure to the +1 stage produced more change than same-stage reasoning with no "stage-skipping" evidenced.

Walker (1982) conducted a study testing the sequentiality assumption using 101 5th to 7th graders. Addressing methodological problems in previous studies, Walker included only subjects that had attained cognitive prerequisites based on Piagetian tasks. In addition, a neutral-treatment group and a no-treatment control group provided greater rigor to the experimental design. The results of the role-playing of reasoning related to the -1, +1, and +2 stages from the subjects dominant stage show no instances of stage-skipping across experimental and control conditions. These results also evidenced stability over time as a 7 week posttest showed no significant changes. Walker attributes the strength of the results supporting sequentiality to changes in the scoring system and to the attainment of prerequisite cognitive skills by the subjects.

Another source of support for sequentiality comes from studies that test for comprehension and preference of moral stages based on Kohlberg's assertion that the stages form an integrated hierarchy of increasing psychological complexity. Subjects' responses should reflect the highest

stage they can fully comprehend. They should also prefer the highest stage they are capable of comprehending.

Rest, Turiel, and Kohlberg (1969) attempted to replicate Turiel's 1966 study as well as determine relevant factors to explain Turiel's results. The authors sought to validate a developmental hierarchy of preference, comprehension, and moral reasoning. They hypothesized that subjects would prefer the increasing complexity of a +1 stage; that subjects could not comprehend and assimilate the reasoning of a +2 stage; and that a lower stage (-1) would be readily reproduced but not judged as adequate. The expectation was that exposure to the +1 stage reasoning would result in maximal change. Experimental conditions involved exposing 5th and 8th graders, matched on gender, to moral dilemmas and associated "advice" as to moral choices. The level of "advice" offered was at the -1, +1, or +2 stage from the subject's pretested level.

The results indicated that the subjects preferred advice from higher levels (+1 & +2) more than their own dominant stage or below (-1). Subjects also had greater difficulty comprehending +2 stages on both individual and group levels. Finally, the subjects were able to assimilate the reasoning behind the +1 more easily than either the -1 or +2 stage reasoning. The findings of this study support the invariant sequentiality as the ascension through the stages show an orderly stepwise progression.

In a similar study Rest (1973) provided more definitive evidence of the hierarchial nature of Kohlberg's stages. The study was based on the hypothesis that subjects would prefer the highest stage that they are able to comprehend. Subjects were given statements for and against making a moral choice and were asked to provide an equivalent recapitulation (paraphrase) of the statement, compare it to their own ideas, and evaluate and rate the statement in terms of relative persuasiveness. The results indicate a relationship between pretest scores and highest level of comprehension, that is, in pretests subjects tended to spontaneously produce, to a limited degree, the highest stage that they were able to comprehend. Subjects also generally preferred stages higher than their own dominant stage. Lower stages were fully comprehended but rejected as inadequate. Half of the subjects were able to comprehend a +1 stage and about 20% demonstrated spontaneous use of this stage in pretests.

The Rest et al. (1969) and Rest (1973) studies support the notion of invariant sequentiality by showing that the stages form a sequence of cumulative difficulty and that preference and comprehension follow predicted patterns related to the concept of hierarchical integration.

An alternative interpretation of these two studies is that they demonstrate a decalage across different domains related to response mode. Subjects may evidence a

particular stage by first expressing a preference on a preference task, then exhibiting comprehension on a paraphrasing-for-comprehension task, and later spontaneously produce stage reasoning on a verbal justification task (Rest, 1979). This interpretation suggests that stage-specific moral reasoning develops progressively based on task-specific response modes. Sequentiality deriving from hierarchial integration assumption may then be more related to the required response mode than a stepwise sequence related to discrete stages.

Additionally, Locke (1979) noted that subjects only use the highest stage they are capable of comprehending 20% of the time. This suggests that factors other than cognitive adequacy operate to produce moral judgment at a given stage.

Evidence Against Invariant Sequentiality

Holstein's (1976) longitudinal study provides the clearest evidence contradicting the assumption of sequentiality. The subjects, 52 adolescents and their parents, were assessed using Kohlberg's moral judgment dilemmas and interviewed at a 3 year interval in an attempt to replicate Kramer's (1968) study.

The results from the two testing times pertinent to the assumption of invariant sequentiality found that adolescent males were likely to move from stage 1 or 2 to stage 4,

indicating a 2 or 3 stage jump. These jumps are comparable to those found in Kramer's 1968 study. Stage-skips of 2 stages, from stage 3 to stage 5, were found in the adolescent and adult subjects. Twenty-one percent of the adolescent sample were found to have skipped stages compared to only 7% of the adult sample suggesting greater stability for adults. Kramer's data on older adolescents indicated a slightly lower figure at 14%.

In general, Holstein found that sequence was maintained if viewed from a level-to-level perspective but not from a stage-to-stage vantage point. Subjects generally showed stepwise movement from the Preconventional level (Stages 1 & 2) to the Conventional level (Stages 3 & 4) but stepwise movement from stage to stage was not supported by the data. The stage-skipping reported in this study contradicts the invariant sequence assumption.

Holstein's study evidences several acknowledged methodological problems that may qualify her findings. First, a three year span may be too long for detecting sequential progression as younger subjects may have progressed through several stages prior to the posttest. Holstein's three year pre-to-posttest interval is comparable to the period used by Kramer (1968) who also found stage jumps similar to Holstein's. However, Colby et al. (1983) also used a three year time period and found no discrepancies with the newer Standard Issue scoring system

suggesting that the lack of stage-skipping may be more related to revisions in the scoring systems than subjects transitioning through intervening stages between testing times.

Secondly, Holstein's design lacked short-term control groups to distinguish individual fluctuations from more stable changes. Thus short-term individual "states" variability may account for the results rather than actual permanent stage changes. These two problems prevent Holstein's (1976) study from representing a clear refutation of invariant sequentiality. However, assuming some degree of state variability exists, this does not account entirely for the discrepancy that Holstein observed. The results of Holstein's study remain equivocal as there has been no effort to replicate her results using a more rigorous design.

Intervention studies have evidenced anomalies related to sequentiality. Arbuthnot (1975), using a role-playing design, found that shifts in moral reasoning scores were greatest when the gap between the subject's stage and the model's stage was widest (+2 stages). His results contradict the research supporting a stage-by-stage ascension through the stages (Rest, 1973; Rest et al., 1969; Turiel, 1966), although the numbers of subjects in the discrepant samples were small.

Another possible understanding of Arbuthnot's (1975) discrepant results is that he used college-age subjects who, having attained a formal operations level of cognitive development, may have had a greater capacity for disequilibrium than the younger subjects used in Turiel's study. Walker's (1982) study, discussed above, used only subjects who had the cognitive prerequisites, as determined by pretests, without finding evidence of stage-skipping. However, Walker's study, like Arbuthnot, also found that using a +2 stage to induce change was just as effective as the +1 condition although the greatest change was found in the +1 stage (stages were not skipped).

Walker (1982) suggested that the +2 stage reasoning was "distorted" toward a +1 stage. This reinterpretation of +2 reasoning would be due to the fact that the subjects had attained the cognitive and perspective-taking prerequisites only for the +1 stage which prevents a full undistorted comprehension of the +2 stage reasoning. Additional research is needed to determine the nature of this anomaly using a larger sample size of subjects who have reached necessary prerequisites for +2 reasoning and exposing them to +1, +2 and +3 stages.

Rest (1983) notes that experimental intervention studies, in general, have not produced "consistent or powerful results" due to problems in assessment and methodology. Problems in isolating causal variables,

unclear definition and support for disequilibrium, lack of sufficiently sensitive measures, minimal treatment effects, and treatments that are too brief to affect the gradual process of moral development limit the support intervention studies can provide for sequentiality.

Conclusions

Overall the empirical evidence for Kohlberg's assumption of sequentiality must be considered equivocal largely due to the use of scoring system revisions to account for stage-skipping in longitudinal studies. Kohlberg's own longitudinal studies provide some fairly strong support for sequentiality but other long-term efforts have found discrepancies that cannot be fully explained by methodological short-comings or measurement error. The lack of well-designed and controlled longitudinal studies is apparent in the empirical literature. Particularly lacking are studies conducted by independent researchers replicating longitudinal studies that are the primary support for invariant sequentiality.

Intervention studies are of limited utility in support of sequentiality due to methodological shortcomings and inconsistent results. Studies focusing on a hierarchial sequence in preference, comprehension, and spontaneous production of moral stage are also of limited value in conclusively demonstrating sequentiality as developmental

decalage across response modes may also explain this hierarchy.

Kohlberg's claims of progressive sequential movement through moral stages have garnered only moderate empirical support. Even if the evidence for an invariant upward sequence were unequivocal such movement would lend more support to a developmental continuum than to Kohlberg's specific stages of moral development. The keystone of Kohlberg's theory is that his stages form discrete wholes and that everyone can be reliably assigned to one of these stages. Kohlberg adopted a "hard" or strict stage theory approach as the basis for his theoretical formulations. For his specific theory of moral development to be distinguished from simple cognitive-developmental age trends, Kohlberg's stages must demonstrate clear internal consistency and "structured wholeness." The evidence for this crucial assumption will be evaluated in the next section.

Single Stage Dominance

Kohlberg's stages are conceptualized as forming a "structured whole" which incorporates the cognitive patterns of each previous stage into a new cohesive organizational schema that is qualitatively distinct from the previous stage (Kohlberg, 1969, 1976). Based on Piaget's notion of "structure d'ensemble" Kohlberg holds to

a strict "hard" stage model based on discrete, holistic, qualitative shifts in the organization of moral reasoning. A person is considered "in" or "out" of a certain stage. Moral stages form a hierarchial integration and use a displacement model of stage change. Disequilibrium, defined in terms of conflict, is the mechanism that provides the impetus or "motor" for upward movement. Each stage should have the property of internal consistency meaning that the individual's moral judgments reflect consistent use of a dominant (modal) stage with limited use of adjacent stages. This stage mixture should represent the process of stage acquisition, that is, transition to the next higher stage.

Single stage dominance is the most crucial issue in an evaluation of Kohlberg's stage theory as it addresses the central concept of a stage theory, that stages are qualitatively distinct entities and demonstrate internal consistency. This section will examine the empirical evidence for single stage dominance.

Evidence Supporting Single Stage Dominance

Single stage dominance predicts that moral judgment scores will show a high degree of internal consistency. The longitudinal data generated by the Colby et al. (1983) study appear to demonstrate such cohesion. Scoring distributions show that most interviews resulted in scores falling in a single dominant stage or one of the two

adjacent stages. The mean percentage of scores in the individual's dominant stage were 67%, 72%, and 69% for forms A, B, and C, respectively. For all forms combined 67% fell in their modal stage. Mean percentages for each form for the two most used adjacent stages were 98% for Form A, 97% for Form B, and 99% for Form C with 99% for all forms combined. These statistics offer strong support for stages as structured wholes and represent a substantial improvement over Kohlberg's (1969) previous reports of 45% scoring at the dominant stage.

Some interviews (9%) earned scores in three adjacent stages. The authors established 10% as a cut-off point where scores that showed stage use below this mark would be treated as measurement error. This choice is derived from the unit of measurement, a "criterion judgment" match, which the authors found to provide a conservative boundary for differentiating actual stage use from error. The found that in every case using 10% and above meant that a full criterion judgment match had been made and stage usage by that individual would exceed the 10% mark.

Additional support from the Colby et al. (1983) study is derived from statistical analyses. A high degree of reliability, alternate form and test-retest, supports the internal consistency criterion. The high level of reliability for this study suggests a consistent cohesiveness that is consonant with the assumption of

internal consistency. However, it is important to note that this level of consistency has not been replicated or found for other cognitive-developmental domains (Flavell, 1971, 1985).

Factor analysis across all dilemmas and issues showed a single factor on which the stage dilemmas and issues loaded heavily also suggesting an underlying stage structure. Colby et al. (1983) asserted that "The absence of issue or dilemma factors along with the absence of scatter across more than two adjacent stages indicates that we have succeeded in defining a coherent moral domain united by a single underlying organizational structure" (p. 73).

Studies by Rest et al. (1969) and Rest (1973) on preference and comprehension, cited earlier as supporting an invariant sequence, also lend support to the assumption of single stage dominance. A subject's comprehension and preference for dominant stage of reasoning and +1 stage reasoning and the rejection of lower stage reasoning point to a certain order and structure. However, discrepancies have been observed in subjects' ability to comprehend a stage outside (+2) of those predicted by a structured whole assumption (Arbuthnot, 1975; Walker, 1982). These problems suggest greater variability within individual development than the concept of single stage dominance could encompass.

Evidence Against Single Stage Dominance

Central to the assumption of stage dominance is the concept of disequilibrium, the mechanism posited by stage theorists for moving from one stage to the next. In Kohlberg's scheme, individuals who evidence "stage mix," that is, score at more than one stage, are thought to be in transition between stages as a result of disequilibrium.

Based on Piaget's (1983) concept of equilibration and described as "cognitive conflict" (Kohlberg, 1971a), disequilibrium is a state of disorganization that allows the integration of the next higher level of moral reasoning by creating transitional "stage mix." For example, one's dominant stage (defined as 50% or more responses coded at that stage) is combined with a secondary stage (20 to 49% of coded responses) resulting in a "mixed" transitional stage that is held to be in a state of disequilibrium. A "pure" stage is thought to be where 50% or greater responses are at a single stage with less than 25% at any other stage, demonstrating less transitional stage mix.

According to Kohlberg (1969), individuals can comprehend their dominant stage and stages below (-1) while the next higher stage (+1) is partially comprehended and preferred creating an imbalance (disequilibrium) that pulls toward upward movement. That same-stage reasoning is preferred over -1 reasoning is well supported by the research (Rest, Turiel, & Kohlberg, 1969; Rothman, 1976).

The literature, however, finds minimal support for the preference of a +1 stage as a means of explaining upward movement. Keasey (1973, 1974) found that subjects preferred same-stage reasoning over +1 reasoning by a slight margin. Rest (1973), using 12th graders, obtained stronger results, finding that Stage 6 was preferred most, followed by Stage 5 and Stage 4.

The conclusion is that while +1 reasoning is preferred over -1 reasoning, preference for a +1 stage over one's dominant stage is not clearly supported. This conclusion casts doubt on Kohlberg's contention that a +1 stage will be preferred, leading to upward change due to a disequilibrated state.

The larger question of whether exposure to +1 reasoning will provide impetus for stage ascension is addressed by Kupfersmid and Wonderly (1982) in a review of the theoretical and empirical support for disequilibrium found mixed results. Several short-term studies indicated little change from the dominant stage with exposure to +1 and -1 stages (Turiel, 1966, 1969; Turiel & Rothman, 1972). Another study found movement at lower stages (Preconventional) and none at higher levels (Tracy & Cross, 1973). Short-term intervention studies do not provide clear evidence of disequilibrium.

Longer-term studies, reviewed by Higgins (1980) and Lockwood (1978), have also found mixed results from

experiments designed to induce upward stage change. These studies used several modes of treatment (verbal, educational, and modeling) to expose subjects to advanced stages of moral reasoning. When change did occur it tended to be minimal with most showing nonsignificant results across conditions.

The method used to induce stage change appears to have an effect on the question of single stage dominance. Arbuthnot (1975), using a role-playing design, found that shifts in moral reasoning scores were greatest when the gap between the subject's stage and the model's stage was widest (+2). In a similar study that also used role-playing to induce change, Keasey (1973) also found that the role-playing intervention at a +1 stage was effective in inducing immediate upward change. However the change obtained did not show stability as a two-week posttest indicated a slight decrease in moral reasoning level. These results counter the assumption of internal structure and suggest that the role-playing method may be a factor in producing change. It may be that role-playing is a more active means of producing the necessary disequilibrium which may allow greater resolution and structural change. More passive methods may not generate sufficient conflict to advance moral reasoning beyond the individual's dominant stage. These results may also reflect the influence of social expectations based on a role-taking opportunity.

Stage change appears to be significantly affected by intervention mode (passive vs. active) suggesting greater variability than a strict stage theory would predict. Additional research is needed where the method of conflict-induction is a primary independent variable.

Long-term studies also provided no support for differentiating "mixed" from "pure" types suggesting that those thought to be in transition (mixed) show no more confusion and conflict on objective tests than those who are "settled" (pure) in their moral judgments (Kupfersmid & Wonderly, 1982). Thus the expected manifestations associated with stage transition were not found.

The conclusion that can be drawn regarding the concept of disequilibrium is that there is little empirical evidence to support its role in stage change. In most cases disequilibrium was either not produced and/or disequilibrium is not a major factor in shifting moral reasoning. The lack of clear evidence for disequilibrium as the mechanism responsible for change in moral reasoning may be due to problems in conceptualization and consequently inadequate operationalization for an experimental test. As Kupfersmid and Wonderly (1982) noted, there is a need for a more adequate operational definition of disequilibrium, greater specificity for disequilibrium-inducing interventions, and a better

understanding of how cognitive, affective, and intuitive components interrelate in producing stage change.

The assumption of stages as forming structured wholes reflect a "hard" concept of stages derived from Piagetian theory and adapted to moral development by Kohlberg. This hard stage concept has been criticized as not conforming to the empirical data. Fischer (1983) is critical of Kohlberg's assertion of stage structure noting that Kohlberg's criteria for structured wholeness reflects a "much weaker" version of the structured whole hypothesis postulated by Piaget (1954). The original Piagetian concept predicts a relatively abrupt and pervasive qualitative change in thinking that evidences a high degree of consistency which is not convincingly demonstrated by the Colby et al. (1983) results.

While it is acknowledged that the statistical support for consistency of stage usage in the Colby et al. (1983) study is greater than other cognitive-developmental domains, Fischer (1983) noted that there is evidence of significant inconsistencies and variability related to environmental influences that counter a "hard" stage theory. For example, one third of the responses on the interviews were outside the subject's modal stage and most subjects showed a range of two or three stages. Different dilemmas and alternate forms of the Moral Judgment Interview resulted in different modal stage assignments

suggesting inconsistency even within the Interview. Further, the stage concept cannot be supported for domains outside the Moral Judgment Interview such as preference and comprehension, dimensions which Kohlberg allows may not show a stage-like distribution (Colby et al., 1983).

Thus the Kohlbergian stage appears to be a relatively narrow concept with strong empirical evidence limited to the domain of the Moral Judgment Interview which Rest (1979) suggested is biased toward internal consistency. Rest alleged that the increased evidence for single stage dominance from revised versions of Kohlberg's scoring system are due to the use of fewer dilemmas (from nine in 1958 to three in 1979), the discarding of "criterion judgments" that produced discrepant scores, and not scoring lower stage responses when they were elaborated at a higher stage later in the interview.

Despite such bias toward consistency Kohlberg's longitudinal study still evidences substantial inconsistencies. These criticisms limit the power of the data from Kohlberg's scoring system in terms of generalizing to other domains and suggest that the degree of stage mixture is underestimated by the Moral Judgment Interview.

Fischer (1983) also observed that, using Kohlberg's data, moral reasoning actually develops slowly and gradually over a substantial period of time. For example,

Stage 4 subjects took 23 years to progress from 5% usage at age 13 to 67% at age 36. This slow movement contradicts the "hard" structured whole concept that predicts a relatively quick transition phase and full stable use of the new stage. Fischer notes that this pattern better fits a model of emergence and generalization where new ways of reasoning gradually spread to other contexts and situations over a period of time. Fischer concluded that a strict stage concept does not fit the data and that environmental situational factors and individual differences need greater attention in Kohlberg's theory.

Related to the fact of gradual acquisition, the issue of decalage is also problematic for Kohlberg's theory. The issue is whether components of a stage of moral reasoning develop synchronistically across the stage or is decalage, defined as a staggered pattern specific to various domains, evidenced across the stage. The significance of decalage is that the individual may have attained and evidence a specific stage in one area or situation but not in another area. That is, the person may "have" a structure but can only manifest such reasoning in certain contexts or on specific tasks. Research has shown considerable support for asynchronous development in cognitive structures (Flavell, 1971). Structures do not develop concomitantly across domains but evidence gradual emergence and generalization. The decalage demonstrated by other

cognitive-developmental domains may also apply to moral development but the lack of well-defined specific components of each stage prevent an investigation into whether these aspects develop synchronously or evidence a "marbling" across ages.

However, given the isomorphism and dependence between cognitive and moral development posited by Kohlberg (1969) the lack of comprehensive discrete stages in cognitive structuralization casts substantial doubt on the existence of content-free synchronous development of moral reasoning. Decalage suggests that moral reasoning will continue to develop through higher stages, generalizing to more and more areas rather than being fully attained and exhibited across all domains as a strict stage theory would require. Studies on preference and comprehension seem to show decalage across tasks (Rest, 1973; Rest, et al., 1969). The assignment to a particular stage based on a specific situation or task (verbal justification on the MJI) will be misleading and unreliable.

Rest (1976, 1979, 1983) was also critical of a "hard" stage concept, alleging that such a rigid and simplistic model is untenable due to individual subject fluctuations, inconsistencies related to test characteristics, lack of synchronous development within stages (decalage), and discrepancies due to response mode (verbal rather than nonverbal). Rest noted that 30% of the subjects in the

Colby et al. (1983) study show short-term fluctuations of one third of a stage over a two week period similar to the variations found by Kuhn (1976). Such inconsistency is significant considering that one third of a stage is the equivalent of four years of natural movement in Kohlberg's longitudinal study. Thus, subjects are not "in" a specific stage but fluctuate within a developmental range of possible moral reasonings depending on contextual factors.

That a strict stage theory such as Kohlberg's is untenable gains significant support from cognitive developmental research on the acquisition of cognitive structures (Flavell, 1970, 1971, 1985; Flavell & Wohlwill, 1969; Wohlwill, 1973). In an extensive review of Piagetian cognitive-developmental literature, structural development was found to be gradual rather than abrupt with significant fluctuations within subjects. Structural attainment was found to be probabilistic in nature; in the short-term structures may be present or not present demonstrating inconsistency rather than an "all-or-nothing" possession as predicted by strict stage theory. These classic studies called into question the viability of comprehensive stage theory and validated the notion that domain-specific content as well as structure operate to determine cognitive development. These findings from the cognitive-developmental literature that Kohlberg relies on to support the extension into the moral domain clearly show that

Kohlberg's attempts to separate content from structure in revisions of the MJI scoring system does not fit the empirical data. The content of moral judgment may be independent of structure to some degree and is relevant in determining moral judgments. That is, both content and structure are essential components of cognitive structures.

Flavell's (1971) paper liberated cognitive-developmental theory to research the nature of domain-specific content and the effects on acquisition of cognitive structures (Damon, 1977; Selman, 1980). Despite the lack of empirical evidence for a simple stage theory, Kohlberg continued to revise his scoring system to obtain a clearer distinction between content and underlying structure.

Based on the substantial inconsistencies in moral stage assignments Rest (1979) contended that a more complex stage theory would better fit the available data than Kohlberg's "simple" stage theory. His alternate approach, using the Defining Issues Test, reflected the view that individuals demonstrate a variety of types of moral reasoning which cannot be legitimately assigned to a single stage. In Rest's model both qualitative (structure) and quantitative (content) analyses are needed to capture the nature of moral judgments. A quantitative analysis of the probability of the use of certain types of reasoning would replace the dichotomous "in" or "out" model of stage

inclusion used by the Kohlbergian system. A person may advance in several organizations (stages) of reasoning simultaneously while showing less use of other types of reasoning. Rest's model reflects a shift in the essential question in moral development research from "What stage is the person in?" to "To what extent and under what conditions does a person manifest the various types of organizations of thinking?" (Rest, 1979, p. 63).

Rest's "types" rather than stages modifies the cognitive-developmental assumption of consistent linear structuralization in favor of a "softer" and "messier" version based on the use of a range of possible cognitive organizations across various contexts. Rest's work appears to move toward a more integrated and flexible model that can incorporate individual and environmental variables into moral reasoning decisions. This model increases the complexity of moral decisions and defines moral development as an intricate, nonlinear process of consolidating ways of reasoning on several different levels simultaneously.

Conclusions

In general, the empirical evidence provides little support for Kohlberg's assertion of a strict stage concept. Despite a moderately high level of consistency using the MJI, which may be related to instrument bias, the longitudinal data exhibit significant inconsistencies that cannot sustain the "structured whole" assumption.

Fluctuations related to environmental and contextual factors, test characteristics, and individual differences do not allow validation of a simple stage model. Research on the acquisition of cognitive structures also demonstrates that a stage theory based on structure alone is not untenable. Kohlberg's continued pursuit and adherence to a strict "hard" version of a stage theory appears to have led his program away from a base of empirical support into a largely unsupported and isolated "hard line" position.

Methodological problems plague research efforts into gaining clearer understanding of the nature of moral structures. Unclear definitions of concepts and components of moral reasoning, few independent replication studies, and partisan allegiance to theoretical positions hamper continued progress in understanding the consistencies and inconsistencies in moral reasoning.

Universality

A primary assumption of Kohlberg's theory is that sequential stepwise progression through moral stages is a universal phenomena occurring in all persons and cultures. Kohlberg maintained that based on universal moral principles "all individuals in all cultures go through the same order or sequences of gross stages of development,

though varying in rate and terminal point of development" (1971a, p. 175).

Cultural Differences

Empirical research on the cross-cultural validity of Kohlberg's stage theory has been conducted in British Honduras (Gorsuch & Barnes, 1973), Turkey (Nisan & Kohlberg, 1982; Turiel, Edwards & Kohlberg, 1978), Israel (Bar Yam, Reimer, & Kohlberg, 1974; Snarey, Reimer, & Kohlberg, 1985), Canada (Kohlberg & Kramer, 1969), Mexico (Kohlberg & Kramer, 1969), Kenya (Edwards, 1978), Great Britain (Kohlberg & Kramer, 1969), Taiwan (Kohlberg & Kramer, 1969), Nigeria (Masqud, 1977, 1979), India (Parikh, 1980), New Zealand (Moir, 1974), and the Bahamas (White, 1975; White, Bushnell, & Regnemer, 1978). This list of studies is not exhaustive; additional sources can be found in reviews by Snarey (1985), Edwards (1986), and Boyes and Walker (1988).

An intensive examination of the cross-national data cannot be undertaken here due to space limitations, however, some general conclusions can be drawn regarding the current status of Kohlberg's cross-cultural universality assumption.

Reviews of the available cross-cultural studies indicate general support for the assumptions of a stage theory (Boyes & Walker, 1988; Snarey, 1985) with some significant qualifications.

Allegations of cultural bias in Kohlberg's theory are frequent in the literature (Simpson, 1974; Vine, 1986). Simpson is critical of Kohlberg for confusing developmental and cultural differences, specifically, for equating cultural differences with a lack of developmental progress. Referring to a less-industrialized culture, Simpson noted that "The need to survive in those particular environments has taught members of these cultures beliefs and values that are not at all likely to be counteracted by developmental processes" (p. 88).

Simpson also points to problems in confusing individual differences with cultural differences, moral development with culturally-determined verbal skills, and culture-specific values of life and property with universal moral givens. The priority of cognitive structure over cultural norms and expectations is also objectional when applied to disparate cultures. (See Kohlberg et al., 1983 for a reply to Simpson.)

Snarey (1985) reviewed 45 cross-cultural moral development studies in terms of their support for Kohlberg's stage theory assumptions. These studies included 7 longitudinal and 38 cross-sectional studies from research in 27 countries. Snarey found "striking" support for Kohlberg's theory with some significant qualifications. Bias toward complex urban societies was identified as more rural cultures lacked significant evidence of principled

moral reasoning. The cross-national data support consistent age differences in moral reasoning through the first three or four stages with higher stages, 4, 5, and 6, rarely appearing in the cross-cultural data.

This suggests that progression to higher stages may be a function of cultural complexity and values (Edwards, 1975; White, 1986; White et al., 1978). Western industrialized societies evidence a more rapid climb through the lower stages with terminal stages in the higher stages more likely than in less developed cultures. Kohlberg's assertion that higher stages are more adequate and "better" than lower stages implies that cultures where upper stages that use autonomous principles to guide moral judgments are not reached are somehow less moral.

White (1986) noted that "All cultures do not emphasize autonomous, self-derived principles as a basis for moral decision-making" (p. 60). The relative absence of upper stages may be explained by incomplete operational definitions of these higher stages, a bias in favor of a Western concept of morality (justice) or a combination of factors. The underrepresentation of upper stages, however, questions the application of Kohlberg's stage theory to non-Western cultures as well as suggesting a greater role for environmental forces than predicted by a stage theory.

Boyes and Walker (1988) also found substantial evidence for Kohlberg's stage theory assumptions in a more recent

review of the cross-cultural data. After examining the cross-cultural evidence for invariant sequence, hierarchial integration, and structural integrity they concluded that "the criteria for a strict moral stage model do hold universally" (p. 51). The authors noted, however, that empirical support does not address the metaethical adequacy of a stage theory. The lack of significant deviations from stage theory assumptions does not assume the cross-cultural validity of a prescriptive application of Kohlberg's theory based on a justice orientation. That is, the empirical support found for Kohlberg's stage theory cannot assume that the totality of the moral domain has been adequately defined and covered for all cultures. Additional work is needed to determine if other modes of moral reasoning better fit the cross-cultural data.

In summary, the cross-cultural evidence appears to support Kohlberg's stage theory based on age trends and lack of significant deviations from stage theory assumptions. However, this evidence does not warrant an uncritical acceptance and application of Kohlberg's theory to all cultures. The influence of cultural norms, the adaptation of the Moral Judgment Interview to diverse cultures, the assumption of justice as the underlying ethical norm, and a relative lack of principled reasoning in some cultures are significant areas in which further study is needed.

Sex Differences

Another difficulty with Kohlberg's assertion of universality is in the "cultural" differences between men and women. Carol Gilligan (1977, 1979, 1982) who collaborated with Kohlberg in several early studies (Kohlberg & Gilligan, 1971) is a leading proponent of sex bias in Kohlberg's scale. Gilligan suggested that Kohlberg's theory does not give adequate consideration to the differences in moral thinking between men and women (Gilligan, 1977, 1982; Gilligan & Belenky, 1980).

Gilligan contended that women speak "in a different voice" of a morality of caring and responsibility over against the predominantly male voice of rights and justice represented by Kohlberg's ideas. Relatedness, empathy, and interpersonal factors characterize a woman's moral reasoning. Gilligan argued that women's responses cannot be integrated into Kohlberg's theory as his concept of morality focuses on male-oriented criteria thereby misrepresenting feminine morality as falling at inferior levels of moral reasoning relative to men. Specifically, women are believed to remain at Stage 3 while most men continue to Stage 4 (Kohlberg & Kramer, 1969).

Gilligan's research focused on the real life decisions of women who are contemplating an abortion. Using Kohlberg's system as a template, interviews with pre-abortion women were conducted, yielding significant

differences in the form and content of moral reasoning (Gilligan, 1977, 1982).

While it is true that Kohlberg and his associates have primarily used men in their studies (Colby et al., 1983), Gilligan's allegation of bias is difficult to sustain as she has done no systematic empirical studies to assess her theory. Her 1977 study used anecdotal interview material to support her theory. Despite the popularity and eloquence of her theory, the contention of sex differences does not stand up to an empirical test at this writing.

Walker (1984) reviewed 79 studies that used Kohlberg's system and found no support for the sex bias hypothesis across ages ranging from childhood to adulthood. Using a meta-analytical methodology, Walker failed to find significant sex differences across the studies reviewed. He noted that the sex differences that were reported may be measurement artifacts of early scoring versions used by the researchers as those using updated scoring forms showed higher levels of moral development for women.

In a mega-review similar to Walker's, Lifton (1985) surveyed 20 different cross-sectional samples totaling over 3,000 subjects. His review determined that sex differences in moral development research are minimal and represent the exception rather than the rule. When differences are found, however, they are usually associated with a cognitive-developmental model such as Kohlberg's. There is

a slight trend favoring girls in childhood and boys in adolescence but the empirical evidence is not strong (Blatt & Kohlberg, 1975; Bussey & Maughan, 1982; Turiel, 1976).

Overall, Lifton's (1985) review did not indicate any consistent differences in favor of either sex. He suggested that the relationship between gender (defined as sex-role socialization) and biological sex is a more fruitful area of study than sex differences alone. Differences may be accounted for by a masculinity/femininity dimension (gender) related to sex-role development rather than biological sex alone.

Rest (1976) noted that Gilligan has moved away from outright sex differences toward an understanding of the care orientation as existing for both men and women, representing one line of moral development. The care orientation along with Kohlberg's justice component describe separate but mutual aspects of moral development which require growth toward higher levels of moral reasoning. Kohlberg's view of this relationship is that "judgments of justice presuppose 'caring' and 'sympathy'; only if the individual sympathizes with the good of others can the justice problem of how the good should be distributed become a problem for moral reasoning" (Colby & Kohlberg, 1987, p. 305).

Kohlberg denied any inherent bias in his theory (Levine, Kohlberg, & Hewer, 1985) stating that "we strongly

disagree with those who make the charge of sexual, cross-cultural, and ideological bias" (Levine et al., 1985, p. 99). Additionally, the authors contend that Kohlberg's theory has been misunderstood and misrepresented while also noting the value of constructive criticism in the ongoing dialogue on moral development.

The substantial empirical evidence does not support Gilligan's theory of sex bias. However, the gender hypothesis (differences on the feminine/masculine dimension) based on socialization factors apart from biological sex is consistent with Gilligan's ideas and warrants further research (Lifton, 1985).

Conclusions

Kohlberg's stage theory of moral development is difficult to evaluate due to several factors. First, frequent revisions of the MJI scoring system prevent direct comparisons of studies using different versions of the scoring system. Each revised method is asserted as more adequate and accurate than its predecessor rendering obsolete prior research employing earlier systems. Although the most recent edition exhibits good psychometric properties, it is difficult to know whether to attribute these improvements to better methodology or a "bootstrapping" circularity that has refined the scoring to fit the theory (Rest, 1979). Additionally, the complexity

of the scoring system and, until recently (1983), its relative inaccessibility to independent researchers, clouds the issue of comparable scoring and methodology in independent tests of Kohlberg's theory.

Second, significant revisions of Kohlberg's theory itself further hampers a clear evaluation. In response to critics and empirical data, Kohlberg has made numerous substantial changes in his theory. The deletion of Stage 6 and the addition of a theoretical Stage 7, the postulation of "hard" and "soft" stages of adult moral development, and the definition of A and B substages to clarify the structure versus content distinction are elaborations of Kohlberg's original stage concept. Reciprocity between critics, data, and theory is developmentally appropriate but also make a focused critique more difficult as these changes, in essence, have redefined Kohlbergian stage theory. This redefinition is particularly evident in the upper stages dealing with adult moral development (Kohlberg, 1973, 1978).

Third, the sheer complexity of the empirical data and analysis presents a formidable task for the critic. Since the validation of Kohlberg's theory rests on available research, not on theoretical debate, the empirical data are the crucial link. Yet exploring these extensive data is often an overwhelming proposition. To negotiate complex statistics, evaluate the methods used, and draw conclusions

regarding the adequacy of the data supporting Kohlberg's theoretical assumptions is a massive undertaking which few attempt (Bergling, 1981).

Despite the inherent difficulties in evaluating Kohlbergian stage theory some conclusions can be formed based on the research reviewed in this paper. There appears to be significant support for upward sequentiality through Kohlberg's stages with several important qualifications. First, the strength of this evidence rests on extensive revisions to the scoring systems which have eliminated deviant findings of regression and stage-skipping. The current scoring system, Standard Issue scoring, exhibits adequate psychometric properties but such properties do not establish its validity for measuring moral reasoning.

Kohlberg's concept of the reciprocal nature of validity and reliability, that empirical support of the stage theory hypotheses also affirm the reliability of the assessment instrument, do not allow confidence that the theorized structural nature of moral stages has been adequately assessed. As Locke (1979) noted, this interdependency makes it "impossible for a valid Kohlbergian measure to refute the invariance thesis!" (p. 173). Such circularity weakens the results of studies reporting support for upward sequentiality. Additionally, Salzstein (1983) observed a parallel between the isolated "cultish" nature of

psychoanalysis and Kohlberg's theory in that "the method may be so tied to the theory that perhaps the latter cannot be used to evaluate the former" (p. 111).

Second, the evidence that does support the upward sequentiality assumption is limited to the lower stages (1-4) of Kohlberg's hierarchy. Kohlberg's research program has encountered substantial problems in supporting the empirical existence of principled Post-Conventional reasoning. The deletion of Stage 6, the extremely limited presence of Stage 5 in Western and cross-cultural studies, and the postulation of "soft" and "hard" stages of moral reasoning strongly suggest that adult moral development cannot be clearly understood in terms of "hard" cognitive-developmental stages (Gibbs, 1977). Adult moral development appears to differ significantly from childhood development that is more closely linked to the emergence of cognitive-logical structures. It appears that, once cognitive prerequisites are attained in adolescence, adult morality must integrate a complex number of individual and situational factors in explaining moral choices.

Another primary conclusion that can be drawn from a review of the research is that Kohlberg's concept of "hard" stages cannot be supported by the available data. The empirical evidence for a strict Piagetian interpretation of moral stages does not hold up under the enormous theoretical weight that Kohlberg places on it. The

presence of significant stage mixture, minimal evidence for disequilibrium, lengthy consolidation periods, the lack of a pervasive shift in reasoning, and decalage across domains make a hard stage concept untenable.

Despite the lack of support for a strict stage interpretation, the cognitive-developmental position is not compromised by the introduction of greater complexity and flexibility in the developmental sequence of moral reasoning. As discussed in the previous section, a more complex model, such as the one proposed by Rest (1976, 1979, 1986) more closely fits the data and is prototypical of the direction of current research trends.

Clinical Implications

The theoretical and clinical question is how the individual's present level of moral development is related to moral decisions, that is, what is the relationship between moral stage and moral behavior? Having a means by which to predict, within limits, the likelihood of a destructive moral decision would be an invaluable tool for clinicians struggling with issues of dangerousness, duty to warn, and social rehabilitation. To date, the empirical research has not provided a clear link between moral thought and moral behavior suggesting a need for continued work toward an understanding of this important relationship.

Directions for Future Research

The direction of cognitive-developmental research on moral development is toward a more complex interactional model, one that can integrate individual differences, personal and situational, into a flexible conceptualization of developmental stages. Rest's (1979) "complex" stage model is prototypical of this trend. Factors such as affect, situational context, individual differences, intelligence, and learning are largely neglected by Kohlberg's stage theory. Recent shifts toward a greater recognition and role for social and individual factors in moral development by cognitive-developmental theorists in general and by those holding a Kohlbergian perspective suggest some recognition of this need (Kurtines & Gerwirtz, 1987). Kohlberg's own theorizing suggested an increasing awareness and movement toward incorporation of social factors. In the most recent formulation of Kohlberg's theory prior to his death, several revisions and additions point to a progression toward a more complex stage theory. These include the introduction of A and B substages to account for the relationship between moral stage and moral action, the postulation of a seventh "soft" stage, the discarding of Stage 6, and increased research on the "socio-moral" atmosphere. However, Kohlberg continued to advance a strict interpretation of a stage theory despite a lack of unequivocal empirical support for such a "hard"

stage theory. Kohlberg's program showed evidence of becoming a "degenerating" research program (Locke, 1979; Phillips & Nicolayev, 1978). Revisions of the scoring systems which have apparently eliminated contradictory anomalies can be seen as "content-decreasing" adjustments designed to protect the "hard core" of stage theory assumptions rather than "content-increasing" scientific explanations.

It appears that it is time for an acceptance of the limitations of a simple stage model and for decisive movement toward a model that can integrate the complex variables that operate in moral decisions. After several decades, Kohlberg's stage model has demonstrated its usefulness and value in stimulating research on moral development. However, substantial doubt remains as to the continued utility in asserting a strict "hard" stage theory. Since his original seminal contribution Kohlberg's research program and theorizing on moral development has been shown to provide a powerful heuristic. His stage theory has been and will continue to be a point-of-departure for moral development research; future efforts will stand on Kohlberg's shoulders. The progressive direction at this juncture is to move beyond Kohlberg toward a more sophisticated complex model that can better explain the multidimensional task of moral reasoning as a distinctly human enterprise.

REFERENCES

Arbuthnot, J. (1975). Modification of moral judgments through role-playing. Developmental Psychology, 11, 319-324.

Bandura, A. (1969). Social learning of moral judgments, Journal of Personality and Social Psychology, 11, 275-279.

Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., & McDonald, F. J. (1963). Influence of social reinforcement and the behavior models in shaping children's moral judgment. Journal of Abnormal and Social Psychology, 67, 274-2(1.

Bar Yam, M., Reimer, J., & Kohlberg, L. (1974). Development of moral reasoning in the kibbutz. Unpublished manuscript, Harvard University.

Bergling, K. (1981). Moral development: The validity of Kohlberg's theory. Stockholm, Sweden: Almquist & Wiksell International.

Blasi, A. (1980). Bridging moral cognition and moral action: A review of the literature. Psychological Bulletin, 88(1), 1-45.

Blatt, M., & Kohlberg, L. (1975). The effects of classroom moral discussion upon children's level of moral judgment. Journal of Moral Education, 4, 129-163.

Boyes, M., & Walker, L. J. (1988). Implications of cultural diversity for the universality claims of Kohlberg's theory of moral reasoning. Human Development, 31, 44-59.

Bussey, K., & Maughn, B. (1982). Gender differences in moral reasoning. Journal of Personality and Social Psychology, 42, 701-706.

Colby, A. (1978). Evolution of a moral development theory. In W. Damon (Ed.), New directions for child development: Moral development (Vol. 2, pp. 89-104). San Francisco: Jossey-Bass.

Colby, A., & Kohlberg, L. (1987). The measurement of moral judgment (Vol. 1). New York: Cambridge University Press.

Colby, A., Kohlberg, L., Fenton, E., Speicher-Dubin, B., & Lieberman, M. (1977). Secondary school moral discussion programmes led by social studies teachers. Journal of Moral Education, 6(2), 90-111.

Colby, A., Kohlberg, L. Gibbs, J., Candee, D., Speicher-Dubin, B., Kauffman, K., Hewer, A., & Power, C. (1983). The measurement of moral judgment: A manual and its results. New York: Cambridge University Press.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral judgment. Monographs for the Society for Research in Child Development, 48(1-2, Serial No. 200).

Conroy, A., & Burton, J. (1980). The trouble with Kohlberg: A critique. Educational Forum, 45, 43-55.

Cowen, P., Langer, J., Heavenrich, J., & Nathanson, M. (1969). Social learning and Piaget's theory of moral development. Journal of Personality & Social Psychology, 11, 261-274.

Damon, W. (1977). The social world of the child. San Francisco: Jossey-Bass.

Edwards, C. P. (1975). Social complexity and moral development: A Kenyan study. Ethos, 3, 505-527.

Edwards, C. P. (1978). Social experience and moral judgment in Kenyan young adults. Journal of Genetic Psychology, 133, 19-29.

Edwards, C. P. (1986). Cross-cultural research on Kohlberg's stages: The basis for consensus. In S. Modgil & C. Modgil (Eds.), Lawrence Kohlberg: Consensus and controversy (pp. 419-430). London: Falmer Press.

Erickson, V. L. (1980). The case study method in the evaluation of developmental programs. In L. Kuhmerker, M. Mentkowski, & V. L. Erickson (Eds.), Evaluating moral development (pp. 151-176). Schenectady, NY: Character Research Press.

Erikson, E. (1963). Childhood and society. New York: Norton.

Evans, C. (1986). Kohlberg's moral judgment interview: Is there a need for additional research? In S. Modgil & C. Modgil (Eds.), Kohlberg: Consensus and controversy (pp. 471-480). Philadelphia: Falmer Press.

Fischer, K. (1983). Illuminating the process of moral development. Commentary in A. Colby, L. Kohlberg, J. Gibbs, & M. Lieberman, A longitudinal study of moral development, Monographs for the Society for Research in Child Development, 48(1-2, Serial No. 200).

Flavell, J. (1970). Concept development. In P. H. Mussen (Ed.), Carmichael's Manual of Child Psychology (3rd ed., Vol. I, pp. 983-1059). New York: Wiley & Son.

Flavell, J. H. (1971). Stage-related properties of cognitive development. Cognitive Psychology, 2, 421-453.

Flavell, J. H. (1985). Cognitive development (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall.

Flavell, J. H., & Wohlwill, J. F. (1969). Formal and functional aspects of cognitive development. In D. Elkind & J. H. Flavell (Eds.). Studies in cognitive development: Essays in honor of Jean Piaget (pp. 69-120). New York: Oxford University Press.

Gibbs, J. (1977). Kohlberg's stages of moral development: A constructive critique, Harvard Educational Review, 47(1), 33-51.

Gilligan, C. (1977). In a different voice: Woman's conception of self and morality. Harvard Educational Review, 47, 481-517.

Gilligan, C. (1979). Woman's place in man's life cycle. Harvard Educational Review, 49, 431-446.

Gilligan, C. (1982). In a different voice: Psychological theory and women's development. Cambridge, MA: Harvard University Press.

Gilligan, C., & Belenky, M. (1980). A naturalistic study of abortion decisions. In R. Selman & R. Yando (Eds.), New directions for child development: Vol. 7. Clinical developmental psychology (pp. 69-90). San Francisco: Jossey-Bass.

Gorsuch, R., & Barnes, M. (1973). Stages of ethical reasoning and moral norms in Carib youths. Journal of Cross-Cultural Psychology, 4, 283-301.

Haan, N. (1982). Can research about morality be scientific? American Psychologist, 37, 1096-1104.

Haan, N., Smith, M. B., & Block, J. (1968). Moral reasoning of young adults: Political-social behavior, family background, and personality correlates. Journal of Personality and Social Psychology, 10, 183-201.

Higgins, A. (1980). Research and measurement issues in moral education interventions. In R. Mosher (Ed.), Moral education: A first generation of research and development (pp. 92-107). New York: Praeger.

Holstein, C. B. (1972). The relation of children's moral judgment level to that of their parents and to communication patterns in the family. In R. C. Smart & M. S. Smart (Eds.), Readings in child development and relationships (pp. 211-220). New York: McMillan.

Holstein, C. B. (1976). Irreversible, stepwise sequence in the development of moral judgment. Child Development, 47, 51-61.

Keasey, C. B. (1973). Experimentally-induced changes in moral opinion and reasoning. Journal of Personality and Social Psychology, 26, 30-38.

Keasey, C. B. (1974). The influence of opinion agreement and quality of supportive reasoning on moral opinion. Journal of Personality and Social Psychology, 30, 477-482.

Kohlberg, L. (1958). The development of modes of moral thinking and choice in the years ten to sixteen. Unpublished doctoral dissertation, University of Chicago.

Kohlberg, L. (1963). The development of children's orientation toward a moral order: Sequence in the development of moral thought. Vita Humanities, 6, 11-33.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), Handbook of socialization theory and research (pp. 347-380). New York: Rand-McNally.

Kohlberg, L. (1971a). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), Cognitive development and epistemology (pp. 445-452). New York: Academic Press.

Kohlberg, L. (1971b). Structural Issue Scoring Manual. Unpublished manuscript, Harvard University, Center for Moral Development and Education, Cambridge, MA.

Kohlberg, L. (1973). Continuities in childhood and adult moral development revisited. In P. Baltes & K. W. Schaie (Eds.), Life-span developmental psychology (pp. 180-210). New York: Academic Press.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), Moral development and behavior (pp. 31-53). New York: Holt.

Kohlberg, L. (1978). Revisions in the theory and practice of moral development. In W. Damon (Ed.), Moral development: New directions for child development (pp. 83-87). San Francisco: Jossey-Bass.

Kohlberg, L. (1981). Essays on moral development. Vol. I. The philosophy of moral development. San Francisco: Harper & Row.

Kohlberg, L. (1984). Essays on moral development. Vol II. The psychology of moral development. San Francisco: Harper & Row.

Kohlberg, L., & Candee, D. (1984). The relation of moral judgment to moral action. In L. Kohlberg (Ed.), Essays on moral development. Vol. II. The psychology of moral development (pp. 498-581). San Francisco: Harper & Row.

Kohlberg, L., Colby, A., Gibbs, J., & Speicher-Dubin, B. (1976). Moral stage scoring manual. Cambridge, MA: Harvard University.

Kohlberg, L., Colby, A., Gibbs, J., Speicher-Dubin, B., & Power, C. (1978). Assessing moral stages: A manual. Unpublished manuscript, Harvard University, Center for Moral Education. Cambridge: MA.

Kohlberg, L., & Gilligan, C. (1971). The adolescent as philosopher: The discovery of the self in the post-conventional world. Daedalus, 100, 1051-1086.

Kohlberg, L., & Kramer, R. (1969). Continuities and discontinuities in childhood and adult moral development. Human Development, 12, 93-120.

Kohlberg, L., Levine, C., & Hewer, A. (1983). Moral stages: A current formulation and a response to critics. New York: S. Karger.

Kramer, R. (1968). Changes in moral judgment response pattern during late adolescence and young adulthood: Retrogression in a developmental sequence. Unpublished doctoral dissertation, University of Chicago.

Kuhn, D. (1976). Short-term longitudinal evidence for the sequentiality of Kohlberg's early stages of moral development. Developmental Psychology, 12(2), 162-166.

Kupfersmid, J., & Wonderly, D. (1982). Disequilibrium as a hypothetical construct in Kohlbergian moral development, Child Study Journal, 12(3), 171-185.

Kurtines, W. M., & Gerwirtz, J. L. (1987). Moral development and social interaction. New York: John Wiley.

Kurtines, W. M., & Greif, E. B. (1974). The development of moral thought: Review and evaluation of Kohlberg's approach. Psychological Bulletin, 81, 453-470.

Kutnick, P. (1986). The relationship of moral judgment and moral action: Kohlberg's theory, criticism, and revision. In S. Modgil & C. Modgil (Eds.), Lawrence Kohlberg: Consensus and controversy (pp. 125-148). Philadelphia: Falmer Press.

Leming, J. S. (1975). An empirical examination of key assumptions underlying the Kohlberg rationale for moral education. Foundational Studies, 3, 3-22.

Leming, J. S. (1983). Foundations of moral education: An annotated bibliography. Westport, CN: Greenwood Press.

Levine, C. (1976). Role-taking standpoint and adolescent usage of Kohlberg's conventional stages of moral reasoning. Journal of Personality and Social Psychology, 34, 41-46.

Levine, C., Kohlberg, L. & Hewer, A. (1985). The current formulation of Kohlberg's theory and a response to critics, Human Development, 28, 94-100.

Lifton, P. (1985). Individual differences in moral development: The relation of sex, gender, and personality to morality. Journal of Personality, 53, 306-333.

Locke, D. (1979). Cognitive stages or developmental stages? A critique of Kohlberg's stage-structural theory of moral reasoning. Journal of Moral Education, 8, 168-181.

Lockwood, A. L. (1978). The effects of values clarification and moral development curricula on school-age children: A critical review of recent research. Review of Educational Research, 48, 325-364.

Masqud, M. (1977). The influences of social heterogeneity and sentimental credibility of moral judgments of Nigerian Muslim adolescents. Journal of Cross-Cultural Psychology, 8, 113-122.

Masqud, M. (1979). Cultural influences on transition in the development of moral reasoning in Nigerian youths. The Journal of Social Psychology, 108, 151-159.

Modgil, S., & Modgil, C. (Eds.). (1986). Lawrence Kohlberg: Consensus and controversy. Philadelphia: Falmer Press.

Moir, D. J. (1974). Egocentrism and the emergence of conventional morality in preadolescent girls. Child Development, 45, 299-304.

Murphy, J., & Gilligan, C. (1980). Moral development in late adolescence and adulthood: A critique and reconstruction of Kohlberg's theory. Human Development, 23, 77-104.

Nisan, M., & Kohlberg, L. (1982). Universality and variations in moral development: A longitudinal and cross-sectional study in Turkey. Child Development, 53, 865-876.

Page, R. A. (1981). Longitudinal evidence for the sequentiality of Kohlberg's stages of moral judgment in adolescent males. The Journal of Genetic Psychology, 139, 3-9.

Parikh, B. (1980). Development of moral judgment and its relation to family environmental factors in Indian and American families. Child Development, 51, 1030-1039.

Peters, R. S. (1971). Morality: A plea for pluralism. In T. Mischel (Ed.), Cognitive development and epistemology (pp. 237-266). New York: Academic Press.

Phillips, D. C., & Nicolayev, J. (1978). Kohlbergian moral development: A progressing or degenerating research program? Educational Theory, 28, 286-301.

Piaget, J. (1954). The construction of reality in the child. New York: Basic Books.

Piaget, J. (1965). The moral judgment of the child. (M. Gabain, Trans.). New York: Free Press. (Original work published 1932).

Piaget, J. (1983). Piaget's theory. In P. H. Mussen (Ed.), Handbook of child psychology (Vol. 1, 4th ed., pp. 103-128). New York: John Wiley.

Prentice, N. M. (1972). The influence of live and symbolic modeling on promoting moral judgment of adolescent delinquents. Journal of Abnormal Psychology, 8, 157-161.

Rest, J. R. (1973). The hierarchial nature of moral judgment: A study of patterns and comprehension and preference of moral stages. Journal of Personality, 41, 86-109.

Rest, J. R. (1976). New approaches in the assessment of moral judgment. In T. Lickona (Ed.), Moral development and behavior (pp. 352-363). New York: Holt, Rinehart & Winston.

Rest, J. R. (1979). Development in judging moral issues. Minneapolis: University of Minnesota Press.

Rest, J. R. (1983). Morality. In P. H. Mussen (Ed.), Handbook of child psychology (4th ed., Vol. 3, pp. 556-629). New York: John Wiley.

Rest, J. R. (1986). Moral development: Advances in research and theory. New York: Praeger.

Rest, J. R., Turiel, E., & Kohlberg, L. (1969). Level of moral development as a determinant of Preference and comprehension of moral judgment made by others. Journal of Personality, 37, 224-252.

Rosenthal, T. L., & Zimmerman, B. J. (1978). Social learning and cognition. New York: Academic Press.

Rothman, G. (1976). The influence of moral reasoning on behavioral choices. Child Development, 47, 397-406.

Rubin, K. H., & Trotter, K. T. (1977). Kohlberg's moral judgment scale: Some methodological considerations. Developmental Psychology, 13, 535-536.

Salzstein, H. (1983). Critical issue in Kohlberg's theory of moral reasoning. Commentary in A. Colby, L. Kohlberg, J. Gibbs, & M. Lieberman, A longitudinal study of moral development, Monographs for the Society for Research in Child Development, 48(1-2, Serial No. 200).

Selman, R. (1980). The growth of interpersonal understanding. New York: Academic Press.

Simpson, E. L. (1974). Moral development research: A case study of scientific cultural bias. Human Development, 17, 81-106.

Snarey, J. (1985). Cross-cultural universality of social-moral development: A critical review of Kohlbergian research. Psychological Bulletin, 97(2), 202-232.

Snarey, J., Reimer, J., & Kohlberg, L. (1985). The sociomoral development of kibbutz adolescents: A longitudinal cross-cultural study. Developmental Psychology, 21, 3-17.

Sullivan, E. V. (1977). A study of Kohlberg's structural theory of moral development: A critique of liberal social science ideology. Human Development, 20, 353-376.

Tracy, J., & Cross, H. (1973). Antecedents of shift in moral judgment. Journal of Personality and Social Psychology, 26, 238-244.

Turiel, E. (1966). An experimental test of the sequentiality of developmental stages in the child's moral judgments. Journal of Personality and Social Psychology, 3, 611-618.

Turiel, E. (1969). Developmental processes in the child's moral thinking. In P. Mussen, J. Langer, & M. Covington (Eds.), Trends and issues in developmental psychology (pp. 92-133). New York: Holt, Rinehart, & Winston.

Turiel, E. (1976). A comparative analysis of moral knowledge and moral judgment in males and females. Journal of Personality, 44, 195-201.

Turiel, E., Edwards, C. P., & Kohlberg, L. (1978). A cross-cultural study of moral development in Turkey and the United States. Journal of Cross-Cultural Psychology, 9, 75-87.

Turiel, E., & Rothman, G. (1972). The influence of reasoning on behavioral choices at different stages of moral development. Child Development, 43, 741-756.

Vine, I. (1986). Moral maturity in socio-cultural perspective: Are Kohlberg's stages universal? In S. Modgil & C. Modgil (Eds.), Lawrence Kohlberg: Consensus and controversy (pp. 431-450). Philadelphia: Falmer Press.

Walker, L. J. (1982). The sequentiality of Kohlberg's stages of moral development. Child Development, 53, 1330-1336.

Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. Child Development, 55, 677-691.

Walker, L. J. (1986). Cognitive processes in moral development. In G. Sapp (Ed.), Handbook of moral development (pp. 109-145). Birmingham, AL: Religious Education Press.

White, C. B. (1975). Moral development in Bahamian school children: A cross-cultural examination of Kohlberg's stages of moral reasoning. Developmental Psychology, 11, 535-536.

White, C. B. (1986). Moral reasoning in Bahamian and U.S. adults and children: A cross-cultural examination of Kohlberg's stages. Behavior Science Review, 20, 47-70.

White, C. B., Bushnell, N., & Regnemer, J. L. (1978). Moral development in Bahamian school children: A three-year examination of Kohlberg's stages of moral development. Developmental Psychology, 14, 58-65.

Wohlwill, J. (1973). The study of behavioral development. New York: Academic Press.